Springer Proceedings in Political Science and International Relations

Carla Henriques
Clara Viseu *Editors*

EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities

The 1st International Conference on Evaluating Challenges in the Implementation of EU Cohesion Policy (EvEUCoP 2022), Coimbra, 2022

OPEN ACCESS

Springer

Springer Proceedings in Political Science and International Relations

Springer Proceedings in Political Science and International Relations brings the most current research presented at conferences and workshops to a global readership. The series features volumes of selected contributions from conferences in all areas of political science, international relations, and public policy. In addition to an overall evaluation by the publisher of the topical interest, scientific quality, and timeliness of each volume, each contribution is refereed to standards comparable to those of leading journals, resulting in authoritative contributions to the respective fields. Springer's production and distribution infrastructure ensures rapid publication and wide circulation of the latest developments in the most compelling and promising areas of research today.

The editorial development of volumes may be managed using Springer's innovative Online Conference Service (OCS), a proven online manuscript management and review system. This system is designed to ensure an efficient timeline for your publication, making Springer Proceedings in Political Science and International Relations the premier series to publish your workshop or conference volume.

Carla Henriques · Clara Viseu
Editors

EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities

The 1st International Conference on Evaluating Challenges in the Implementation of EU Cohesion Policy (EvEUCoP 2022), Coimbra, 2022

Springer

Editors
Carla Henriques
Coimbra Business School
Polytechnic Institute of Coimbra
Coimbra, Portugal

Clara Viseu
Coimbra Business School
Polytechnic Institute of Coimbra
Coimbra, Portugal

ISSN 2731-6890 ISSN 2731-6904 (electronic)
Springer Proceedings in Political Science and International Relations
ISBN 978-3-031-18160-3 ISBN 978-3-031-18161-0 (eBook)
https://doi.org/10.1007/978-3-031-18161-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Organization

Organizing Committee

Carla Henriques (Chair)
Ana Cristina Amaro
António Trigo Ribeiro
Clara Pisco Viseu
Maria do Castelo Gouveia

Scientific Committee

Ana Camanho (Faculdade de Engenharia da Universidade do Porto)
Ana Cristina Amaro (Polytechnic of Coimbra, Coimbra Business School | ISCAC and CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Portugal)
António Trigo Ribeiro (Polytechnic of Coimbra, Coimbra Business School | ISCAC and Centro Algoritmi)
Carla Oliveira Henriques (Polytechnic of Coimbra, Coimbra Business School | ISCAC, INESC Coimbra and CeBER)
Clara Pisco Viseu (Polytechnic of Coimbra, Coimbra Business School | ISCAC and CMA-UBI)
Carlos Santos Silva (IN+, Instituto Superior Técnico, Universidade de Lisboa, Portugal)
Gerhard-Wilhelm Weber (Poznan University of Technology)
Chiara D'Alpaos (Dipartimento ICEA, Università degli Studi di Padova)
José Rui Figueira (CEG-IST, Instituto Superior Técnico, Universidade de Lisboa, Portugal)
Luis Lopez-Agudo (Department of Applied Economics, University of Malaga)
Luis Dias (University Coimbra, CeBER, Faculty of Economics)
Maria Barbati (Faculty of Business and Law, University of Portsmouth)

Maria da Conceição Andrade Silva (Portuguese Catholic University)
Maria do Castelo Gouveia (Polytechnic of Coimbra, Coimbra Business School | ISCAC, INESC Coimbra and CeBER)
Maria Elisabete Duarte Neves (Polytechnic of Coimbra, Coimbra Business School | ISCAC and INESC Coimbra)
Mónica Domínguez Serrano (Universidad Pablo de Olavide, Sevilla)
Milagros Baldemor (Don Mariano Marcos Memorial State University, South La Union Campus)
Nina Kajiji (Computer Science and Statistics, University of Rhode Island)
Oscar Marcenaro (Department of Applied Economics, University of Malaga)
Patrícia Pereira da Silva (University Coimbra, CeBER, Faculty of Economics)
Pedro Cerqueira (University Coimbra, CeBER, Faculty of Economics)
Salvatore Corrente (Department of Economics and Business, University of Catania)
Salvatore Greco (Department of Economics and Business, University of Catania, and Portsmouth Business School)
Sara María González Betancor (Department of Quantitative Methods, Universidad de Las Palmas de Gran Canaria)
Sérgio Pereira dos Santos (Faculty of Economics, University of Algarve)
Teresa Cardoso Grilo (Department of Marketing, Operations and General Management ISCTE Business School)

Preface

The European Union (EU) cohesion policy aims at attaining the convergence of economic, social, and territorial cohesion across EU Member States (MS). The funds dedicated to the EU cohesion policy constitute the second-highest group of expenditures in the EU budget. Therefore, the evaluation of their implementation assumes a prominent role in cohesion policy formulation because it helps support policy design, and also provides sound evidence of the results and impacts of the actions undertaken. During the 2014–2020 programming period, MS became obligated (for the first time) to conduct evaluations to appraise the effectiveness, efficiency, and impact of each programme's objective. These rules apply to the European Regional Development Fund (ERDF), the European Social Fund, and the Cohesion Fund.

Most of these assessments are focused on implementation matters and evaluate progress regarding targets achievement, being mainly concerned with the alignment of the projects and actions with the programmes' objectives as well as with the effectiveness and efficiency of their implementation. These evaluations also focus on whether the existing funding is spent or not and if the targets established, particularly those of the performance framework, are achieved. The impact assessments are performed later in the programme cycle when most actions already took place and have also generated impacts.

This Book presents recent findings, sparks discussion, and reveals new research paths addressing the use of novel methodologies and approaches to tackle the challenges and opportunities that are unveiled with the implementation of EU cohesion policy. The authors cover a wide range of topics including (but not limited to) the monitoring data; the clearness of indicators in measuring the impact of interventions;

evaluation methods; case studies and applications on evaluations of the thematic objectives under scrutiny of the cohesion policy, namely:

- Research, Technological Development and Innovation;
- Information and Communication Technologies;
- Shift towards a low-carbon economy.

Carla Henriques
Polytechnic of Coimbra
Coimbra Business School | ISCAC,
Quinta Agrícola
Bencanta, Coimbra, Portugal

INESC Coimbra – DEEC
University of Coimbra
Polo, Coimbra, Portugal

Univ Coimbra, CeBER
Faculty of Economics
Av Dias da Silva, Coimbra, Portugal

Clara Viseu
Polytechnic of Coimbra
Coimbra Business School | ISCAC,
Quinta Agrícola
Bencanta, Coimbra, Portugal

CMA-UBI
Universidade da Beira Interior
Covilhã, Portugal

Sponsors and Institutional Support

This book was financed by:

The workshop took place as part **of a scientific research project that aims to develop** an innovative methodology **to support the monitoring and evaluation of the implementation of programmes around three major thematic axes: Research and Innovation; Low carbon economy; and Information & Communication Technologies**. This project is co-financed by the European Regional Development Fund (ERDF) through the Portugal 2020 Partnership Agreement—Technical Assistance Operational Programme (POAT 2020) under the POAT-01-6177-ERDF-000044 ADEPT: Evaluation of Co-Financed Intervention Policies in Enterprises.

Organizing Institutions

Coimbra Business School | ISCAC

Institute for Systems Engineering and Computers at Coimbra

Contents

Mid-term/terminal Assessment: Data Envelopment Analysis Approaches

Evaluating the Reasons Behind the Inefficient Implementation of ERDF Devoted to R&I in SMEs

Carla Henriques and Clara Viseu

Abstract This work is mainly aimed at evaluating the reasons behind the inefficient execution of Operational Programs (OPs) aimed at promoting research and innovation (R&I), especially in small and medium-sized enterprises (SMEs). To achieve this goal, we employed a three-stage slack-based measure (SBM) data envelopment analysis (DEA) model combined with Stochastic Frontier analysis (SFA), which includes a multiplicity of achievement metrics and environmental factors, to evaluate 53 OPs from 19 countries. Our findings suggest that more developed regions (proxied by a higher Gross Domestic Product (GDP) per capita) do not make an efficient application of European Regional Development Funds (ERDF) aimed at fostering R&I in SMEs. Also, a greater proportion of the population with a university degree does not imply an appropriate use of ERDF devoted to R&I in SMEs. Lifelong learning is positively linked with the performance of the outcomes "Researchers Working in Improved Infrastructures" and "Enterprises Supported". Research and development (R&D) expenditures in the public sector contribute favorably to the needed improvements in "Researchers Working in Improved Infrastructures" but have the reverse effect on the number of "Enterprises Supported" and "Enterprises Working with Research Institutions". Furthermore, because R&D expenditures in the business sector have a positive impact on the necessary development of "Enterprises Working with Research Institutions", these results appear to demonstrate that public R&D has a weaker influence on SME innovation than private R&D. Finally, innovative SMEs collaborating with other sources of knowledge show a positive effect on both

C. Henriques (✉) · C. Viseu
Polytechnic of Coimbra, Coimbra Business School, 3045-601 Coimbra, Portugal
e-mail: chenriques@iscac.pt

C. Henriques
INESC Coimbra—Department of Electrical and Computer Engineering (DEEC), University of Coimbra, Polo 2, 3030-290 Coimbra, Portugal

Centre for Business and Economics Research (CeBER), Faculty of Economics, University of Coimbra, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

C. Viseu
Centro de Matemática e Aplicações (CMA-UBI), Universidade da Beira Interior, 6200-001 Covilhã, Portugal

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_1

the number of "Enterprises" and "Enterprises Working with Research Institutions" supported.

Keywords R&I · SMEs · SBM model · SFA · ERDF

1 Introduction

When it comes to innovation, SMEs have a variety of practical challenges. Accessibility to finance may be challenging to get for SMEs, particularly when risky initiatives are involved (Lee et al., 2010; Romero-Martínez et al., 2010; Van de Vrande et al., 2009). The level to which this is an obstacle differs depending on the age of the organization, the company size, the intensity of the investigation, the growth orientation (Zimmermann & Thomä, 2016), and, in many circumstances, the geographic location (Hölzl & Janger, 2014). Additional hurdles may include problems in hiring highly trained individuals (Belitz & Lejpras, 2016; Bianchi et al., 2010; Dahlander & Gann, 2010; Duarte et al., 2017; Gardocka-Jałowiec & Wierzbicka, 2019), issues with management (Zhou et al., 2021), lower adsorption ability (Müller et al., 2021), and challenges in capturing value (Bouncken et al., 2020). Nonetheless, the most fundamental hurdles to innovation are perceived to be economic (García-Quevedo et al., 2018). Over the 2014–2020 programmatic period, the ERDF provided around 66 billion Euros to boost innovation and productivity, particularly in the European Union (EU) SMEs (Gramillano et al., 2018). Despite the evaluation of the implementation of these funds being mandatory, according to Ortiz and Fernandez (2022), policymakers still face major challenges in their assessment and control stages, owing to the absence of useful information, comparative studies, and organizational qualifications. Moreover, evaluation mechanisms during the 2014–2020 programmatic cycle focused heavily on evaluating procedure results, with hardly any data on the criteria to measure the immediate benefits of the initiatives funded (Ortiz & Fernandez, 2022). Furthermore, in the case of R&I policies, the assessment technique plays an important role in assisting the national/regional authorities in the enhancement of upcoming policy tools by identifying the strengths and weaknesses of previous policy stages (Neto & Santos, 2020). In this context, there are numerous techniques for appraising cohesion policy (Lopez-Rodríguez & Faíña, 2014). Macroeconomic and econometric modeling are commonly used approaches for analyzing the effect of cohesion policy (Henriques et al., 2022a, b). Computable General Equilibrium Models along with input–output models and econometric techniques are normally employed in the context of R&I socioeconomic effect evaluation (e.g., Di Comite et al., 2018; Diukanova et al., 2022; Barbero et al., 2022). Even though these approaches allow for the evaluation and study of the major effects of EU funds on economic growth, they do not allow evaluating management failures (Marzinotto, 2012). Moreover, they ignore the allocation of EU funding within every region to different thematic objectives (TO). The research mainstream is based on econometric studies (see, for example, Stojčić et al., 2020; Radicic & Pugh, 2017; Santos et al., 2019; Thum-Thysen et al., 2019; Fattorini et al., 2020; Sein & Prokop, 2021). Nevertheless, it produces contradictory results (Berkowitz et al., 2019), prompting some experts to dispute its use (Durlauf, 2009; Wostner & Šlander, 2009; Berkowitz et al., 2019). Other methods can also be employed, but with the same intrinsic shortcomings (e.g., Bedu & Vanderstocken, 2020; Gustafsson et al., 2020). The evaluation procedures generally available do not allow comparing any regional or national OP against its peers. These do not enable the identification of the adjustments that should occur to enhance the efficiency of OPs' execution (Gouveia et al., 2021). Moreover, these methods often require fulfilling statistical hypotheses (namely, normality, absence of multicollinearity, and homoscedasticity). Therefore, the adoption of nonparametric methodologies can be valuable and appropriate, particularly as the data freely available on the European Commission website can be used in conjunction with DEA models. The efficient production frontier is usually derived through stochastic approaches (Gouveia et al., 2021). These, nevertheless, can just accommodate an output level at a time (Gouveia et al., 2021). Contrastingly, DEA can easily handle many inputs (resources) and outputs (outcomes) and can also be applied to determine the efficient production frontier. Furthermore, contrary to stochastic techniques, DEA does not rely on any production function form or error term. According to DEA, the greater the divergence from the production efficient frontier, the greater the inefficiency of the decision-making unit (DMU) (in this case, the OPs) under appraisal. Also, the DEA methodology can be particularly valuable for management authorities (MA) because it enables the detection of best practices, and also identifies the changes that need to occur to improve the performance of the OPs under evaluation.

In this framework, Athanassopoulos (1996) used DEA to determine the relative geographical weaknesses of the EU's Level II territories. Gómez-García et al. (2012) assessed the pure and global technical efficiencies regarding Thematic Objective 1 (TO1) in the deployment of EU structural funds from 2000 to 2006. They employed labor and productivity levels as outputs, and the Stochastic Frontier Analysis (SFA) together with the DEA methodology. Anderson & Stejskal, (2019) employed DEA to evaluate the efficiency of innovation diffusion in EU MS based on their European Innovation Scoreboard scores. Furthermore, Gouveia et al. (2021) employed the Value-Based DEA technique, considering the primary elements that can impair the efficient execution of structural funds in different OPs devoted to SMEs' competitiveness. Henriques et al. (2022a) used the SBM approach in conjunction with cluster analysis to evaluate 102 OPs from 22 EU MS focused on the implementation of a low-carbon economy in SMEs. Finally, Henriques et al. (2022b) evaluated the efficiency of 53 R&I OPs from 19 countries utilizing the Network SBM technique in combination with cluster analysis for appraising the implementation of EU funds devoted to promoting R&I in SMEs. Nevertheless, their work did not accommodate for the influence of contextual variables and random errors in efficiency evaluation. Therefore, this work aims to fill this gap by suggesting an approach that combines a three-stage SBM model and SFA, which to the best of our knowledge has not hitherto been used in this context. Through this method it is possible to further understand if the efficiency results attained are mainly related to management failures or the contextual environment of the OPs or statistical noise, also providing information on the contextual factors with the greatest effect on the OPs' inefficiencies.

Insofar, the main research questions that we seek to address with this work are given below:

RQ1: "Which contextual variables show a relevant effect on the inefficiencies of the OPs committed to boosting R&I in SMEs?"
RQ2: "What are the impacts of considering contextual factors on the efficiency of the OPs?"

This article is organized as follows. Section 2 explains the basic assumptions underlying the techniques suggested to assess the execution of the OPs evaluated. Section 3 addresses the key rationale for choosing the inputs and outputs utilized in this study, as well as some statistics on the data that instantiates the SBM and SFA models. Section 4 delves into the major findings. Section 5 summarizes the major results, discusses potential policy recommendations, identifies the main shortcomings, and proposes further work advances.

2 Methodology

Classical DEA techniques, like the CCR (Charnes et al., 1978) and BCC (Banker et al., 1984), are radial, which means that they can simply manage proportional adjustments in the inputs or outputs used in the assessment. Therefore, the CCR and BCC efficiency ratings produced indicate the highest proportionate input (output) contraction (expansion) rates for all inputs (outputs). Nevertheless, owing to factor substitutions, this sort of premise is frequently not met in practice.

As a result, in opposition to the CCR and BCC approaches, we employ the SBM approach (Tone, 2001), which allows for a broader study of efficiency due to its non-radial nature (i.e., inputs and outputs can vary non-radially), also enabling to consider non-oriented models (i.e., address simultaneous variations of the inputs and outputs).

2.1 The SBM Model

The generalized SBM model of Tone (2001) may be presented (by taking m inputs, s outputs, and n DMUs into account) as follows:

$$
\begin{aligned}
&\underset{\lambda,\mathbf{s}^-,\mathbf{s}^+}{Min}\ \rho = \frac{1-\frac{1}{m}\sum_{i=1}^{m} s_i^-/x_{ik}}{1+\frac{1}{s}\sum_{r=1}^{s} s_r^+/y_{rk}} \\
&\text{s.t.} \\
&x_{ik} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, i = 1, ..., m \\
&y_{rk} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, r = 1, ..., s \\
&\sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \ \ldots, n, \\
&\lambda_j \geq 0, j = 1, ..., n, \\
&s_i^- \geq 0, i = 1, ..., m, \\
&s_i^+ \geq 0, r = 1, ..., s,
\end{aligned}
\tag{1}
$$

where $X = [x_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n]$ is the $(m \times n)$ matrix of *inputs*, $Y = [y_{rj}, r = 1, 2, \ldots, s, j = 1, 2, \ldots, n]$ is the matrix of *outputs* $(s \times n)$ and the rows of these matrices for DMU_k are, respectively, $\mathbf{x}_k^T$ and $\mathbf{y}_k^T$, where $^{\text{T}}$ is the transpose of a vector. Also, we presume a Variable Returns to Scale technology with the imposition of $\sum_{j=1}^{n} \lambda_j = 1, {}_j \geq 0\ (\forall_j)$. The value of $0 < \rho < 1$ can be seen as the ratio of average inefficiencies of inputs and outputs.

A DMU_k is SBM-efficient if $\rho^* = 1$, , meaning that the slacks (s_i^- and s_i^+) are null for all the inputs and outputs.

Problem (1) can be converted into a linear problem, by applying a positive scalar variable t (see Tone (2001)). Further details on this modeling approach can be found in Tone (2001) and regarding SBM superefficiency in Tone (2002).

2.2 *Stochastic Frontier Analysis*

Fried et al. (2002) proposed a three-stage DEA model. In the first stage, the SBM model is applied to calculate the technical efficiency of each DMU, and the necessary changes required to the inputs and outputs to turn inefficient DMUs into efficient ones (i.e., the slacks). In the second stage, the slacks are grouped into three types: contextual variables, inefficient management, and statistical noise. The slacks are the dependent variables, while the contextual variables are the independent variables. The objective is to remove the influence of contextual factors and random errors. SFA is then used to modify the input and output factors (Aigner et al., 1977; Meeusen & Broeck, 1977).

Therefore, the slack of each input obtained for every inefficient DMUj ($j = 1, \ldots, p$) is:

$$s_{ij} = f\left(X_j, \beta^i\right) + v_{ij} + u_{ij}, i = 1, \ldots, m; j = 1, \ldots, p, \tag{2}$$

where s_{ij} is the slack of input i of DMUj, $f\left(X_j, \beta^i\right)$ is the slack frontier, and β^i corresponds to the coefficients related to the contextual variables. Expression $v_{ij}+u_{ij}$ is the mixed error, v_{ij} is the statistical noise and u_{ij} is the management inefficiency. Generally, it is presumed that $v_{ij} \sim N\left(0;\ \sigma_v^2\right)$ and $u_{ij} \sim N^+\left(\mu^i; \sigma_u^2\right)$, where v_{ij} and u_{ij} are independent variables.

Consider that $\gamma = \frac{\sigma_u^2}{\sigma_u^2+\sigma_v^2}$. If γ is near 1, it implies that the majority of the adjustment necessary to reach efficiency is related to management inefficiency. If γ is near 0, the random error is the prevalent factor.

Subsequently, the adjusted input and output slacks are obtained by splitting the mixed error. According to Jondrow et al. (1982), the conditional inefficiency is given as:

$$E\left(u_{ij}|u_{ij} + v_{ij}\right) = \frac{\sigma\delta}{1+\delta^2}\left[\frac{\varphi\left(\frac{\varepsilon_j\delta}{\sigma}\right)}{\emptyset\left(\frac{\varepsilon_j\delta}{\sigma}\right)} + \frac{\varepsilon_j\delta}{\sigma}\right], \tag{3}$$

where $\delta = \frac{\sigma_u}{\sigma_v}, \varepsilon_j = v_{ij} + u_{ij}, \sigma^2 = \sigma_u^2 + \sigma_v^2, \varphi$ and $\emptyset$ are, correspondingly, the density and distribution functions of the standard normal distribution. Hence, the expected value of random error is:

$$E\left(v_{ij}|u_{ij} + v_{ij}\right) = s_{ij} - f\left(Z_j, \beta^i\right) - E\left(u_{ij}|u_{ij} + v_{ij}\right), \tag{4}$$

Secondly, the input and output factors of each DMU are changed according to the SFA outcomes by removing the significant contextual effects and statistical noises.

According to Tone and Tsutsui (2009), we begin by employing these formulas:

$$x_{ij}^A = x_{ij} - f\left(Z_j, \hat{\beta}^i\right) - \hat{v}_{ij}(input) \tag{5}$$

$$y_{rj}^A = y_{rj} + f\left(Z_j, \hat{\beta}^r\right) + \hat{v}_{rj}(output). \tag{6}$$

The input data are adjusted using (5) as follows (Tone & Tsutsui, 2009):

$$x_{ij}^{AA} = \frac{x_{imax} - x_{imin}}{x_{imax}^A - x_{imin}^A}\left(x_{ij}^A - x_{imin}^A\right) + x_{imin}, i = 1, \ldots, m; j = 1, \ldots, p \tag{7}$$

where

$$x_{imin} = \min_k\{x_{ik}\}; x_{imax} = \max_k\{x_{ik}\}; x_{imin}^A = \min_k\left\{x_{ik}^A\right\} \text{ and } x_{imax}^A = \max_k\left\{x_{ik}^A\right\}.$$

Analogously, the outputs are changed using (6) as (Tone & Tsutsui, 2009):

$$y_{rj}^{AA} = \frac{y_{rmax} - y_{rmin}}{y_{rmax}^{A} - y_{rmin}^{A}}\left(y_{rj}^{A} - y_{rmin}^{A}\right) + y_{rmin}, r = 1, \ldots, s; j = 1, \ldots, p \quad (8)$$

where

$$y_{rmin} = \min_{k}\{y_{rk}\};\ y_{rmax} = \max_{k}\{y_{rk}\};\ y_{rmin}^{A} = \min_{k}\left\{y_{rk}^{A}\right\} \text{ and } y_{rmax}^{A} = \max_{k}\left\{y_{rk}^{A}\right\}.$$

Then again, the efficiency scores are computed through SBM by employing the previously adjusted inputs and outputs.

3 Data

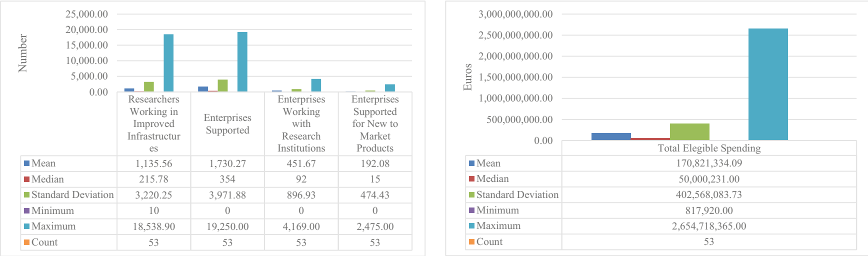
3.1 Input and Output Factors

This work is a follow-up of the work published by Henriques et al. (2022b) and, therefore, we have employed mostly the same input and output factors chosen therein for evaluating the efficiency of the execution of ERDF allotted to boost R&I in SMEs—see Table 1 and Fig. 1. All the information regarding these data is obtainable from Henriques et al. (2022b).

Table 1 External and intermediate inputs and outputs selected for instantiating the SBM model

	Researchers working in improved infrastructures	Enterprises supported	Enterprises working with research institutions	Enterprises supported for new to market products	Total eligible spending
Description	Number of researchers working in improved research infrastructures	Number of enterprises supported	Number of enterprises cooperating with research institution	Number of enterprises supported to introduce new-to-the-market products	Eligible costs validated
Type of factor	Output	Output	Output	Output	Input
Unit	Number of researchers full time equivalent	Number of enterprises	Number of enterprises	Number of enterprises	Euro
Classification	Output indicator	Process indicator	Output indicator	Process indicator	Financial indicator

Source Based on Henriques et al. (2022b)



	Researchers Working in Improved Infrastructures	Enterprises Supported	Enterprises Working with Research Institutions	Enterprises Supported for New to Market Products
Mean	1,135.56	1,730.27	451.67	192.08
Median	215.78	354	92	15
Standard Deviation	3,220.25	3,971.88	896.93	474.43
Minimum	10	0	0	0
Maximum	18,538.90	19,250.00	4,169.00	2,475.00
Count	53	53	53	53

(a)

	Total Elegible Spending
Mean	170,821,334.09
Median	50,000,231.00
Standard Deviation	402,568,083.73
Minimum	817,920.00
Maximum	2,654,718,365.00
Count	53

(b)

Fig. 1 Descriptive statistics of Inputs and Outputs for output and process indicators **a** and for financial indicators **b**. *Source* Authors' computation based on data from Henriques et al. (2022a, b)

3.2 Contextual Factors

The regional GDP at purchasing power parity per capita (GDPPPpc) was considered a contextual variable being used as a proxy to measure economic activity (Barbero et al., 2022; Diukanova et al., 2022; Hervás-Oliver et al., 2021). Besides, Barbero et al. (2022) concluded that the achievement of regional targets related to the ERDF TO1 has a positive impact on all economic indicators, including the GDP, in the selected regions (Greece, Italy, Portugal, and Spain).

According to Diukanova et al. (2022), R&I and low-carbon European structural funds can exert substantial positive effects on the indicator of tertiary education attainment, thus the percentage of the population aged 25–34 who have finished university education was also considered in this set of contextual factors.

Anderson and Stejskal (2019) used variables that fall into the category of human resource (lifelong learning, employment in knowledge-intensive activities), finance (public sector R&D expenditure, private sector R&D expenditure, sales of new-to-market and new-to-firm innovations) and non-financial innovation structures (non-R&D innovation expenditure). Additionally, as referred in Hervás-Oliver et al. (2021), the variation in the development of EU regions affects the innovation capacity of SMEs located in each territory and consequently, it is important to incorporate the variables that better capture innovation in SMEs (e.g., innovation activities like public and private R&D expenditures, non-R&D innovation expenditures, innovative SMEs collaborating with others).

Therefore, we have used similar variables that were reported in the latest European Innovation Scoreboard (Hollanders, 2021). Finally, Sein and Prokop (2021) stress the key role of a firm's R&D, which has proven to be a mediator of the effects of public funding and triple- and quadruple-helix cooperation on the product and process innovation activities of Norwegian firms. In this study, variables such as SMEs with product innovations and SMEs with business process innovations were used. Therefore, we considered, in this context, sales of new-to-market and new-to-firm innovation. All the contextual variables shown in Table 2 (apart from GDPPPPpc whose data were obtained from the OECD website) were extracted from the European

Table 2 Descriptive statistics of the contextual variables

Contextual variables	Mean	Standard deviation	Min	Max
$GDPPPP_{PC}$	92.39	36.52	49.09	269.40
Population with tertiary education	0.4807	0.2513	0.0512	1
Lifelong learning	0.3649	0.2075	0.0145	1
R&D expenditures public sector	0.4568	0.2386	0.0225	1
R&D expenditures business sector	0.2781	0.2211	0.0143	1
Non-R&D innovation expenditures	0.4573	0.2506	0	1
Innovation expenditures per person employed	0.5056	0.2039	0.0449	1
Product process innovators	0.5519	0.2602	0.0460	1
Business process innovations	0.5822	0.3139	0	1
Innovative SMEs collaborating with others	0.4401	0.2119	0.0566	1
Patent Cooperation Treaty (PCT) patent applications	0.3874	0.2592	0	1
Employment knowledge-intensive activities	0.4309	0.2346	0.0071	1
Employment in innovative SMEs	0.5576	0.3056	0	0.9944
Sales of new-to-market and new-to-firm innovations	0.5150	0.1837	0.1148	0.8084

Source Authors' own elaboration

Innovation Scoreboard (Hollanders, 2021), allowing to capture differences in SMEs innovation across regions. All these indicators are normalized between 0 and 1 at origin, to produce a composite indicator integrating variables from different scales. Table 2 shows the main descriptive statistics of the contextual variables.

4 Discussion of Results

The initial results were computed with the help of the Max DEA software and their descriptive statistics are depicted in Table 3.

From Table 3, it can be seen, in general, that the variability of the efficiency scores is bigger for efficient OPs than for inefficient ones (with the standard deviation varying between 0.25 and 0.15, for the first and the latter, respectively). Besides, inefficient OPs present very low mean efficiency scores (with an average potential improvement of efficiency of 94%). Figure 2 illustrates the number of OPs at several subintervals for the efficiency scores.

The number of OPs classified as efficient is 10 (Fig. 2).

Table 3 Descriptive statistics for efficient and inefficient OPs

	Statistics	Efficiency score	Researchers working in improved infrastructures	Enterprises supported	Enterprises working with research institutions	Enterprises supported for new to market products	Total eligible spending
Efficient OPs	Mean	1.32	3,917.39	5,436.94	1,109.54	498.40	227,352,595.60
	Median	1.24	336.50	1,545.00	647.00	12.00	44,310,411.50
	Standard deviation	0.27	6,870.73	7,597.35	1,438.08	994.60	405,454,176.09
	Minimum	1.04	77.68	0.00	0.00	0.00	817,920.00
	Maximum	1.88	18,538.90	19,250.00	4,169.00	2,475.00	1,278,171,878.00
	Count	10	10	10	10	10	10
Inefficient OPs	Mean	0.06	488.62	868.26	298.68	120.84	157,674,529.09
	Median	0.01	161.28	330.00	91.00	15.00	50,000,231.00
	Standard deviation	0.15	672.51	1,769.60	652.57	197.88	405,563,908.88
	Minimum	0.00	10.00	0.00	0.00	0.00	1,424,378.00
	Maximum	0.88	2,840.20	9,677.00	3,590.18	660.00	2,654,718,365.00
	Count	43	43	43	43	43	43

Source Authors' own elaboration

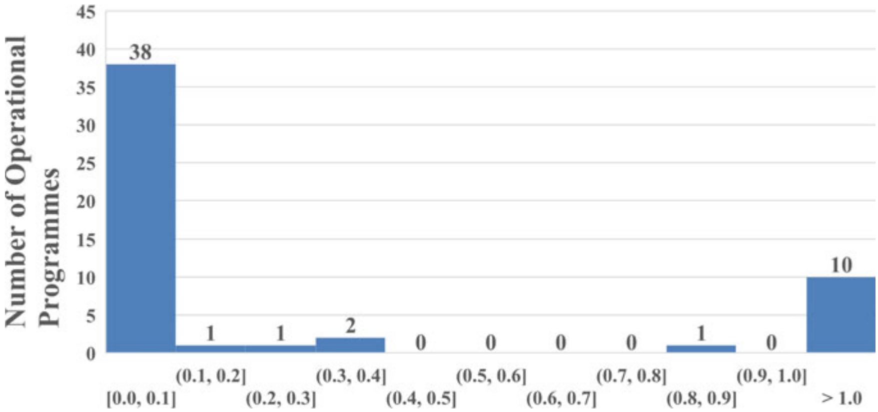


Fig. 2 Number of OPs at different subintervals of efficiency scores. *Source* Authors' own elaboration

Out of the ten efficient OPs, the three most chosen as benchmarks are "Brussels Capital Region—ERDF" (41 times), "Aragón—ERDF" (17 times), and "Toscana—ERDF" (18 times)—see Table 4. The OP most frequently viewed as a benchmark is characterized as an "Innovation Leader" and manages to score in all the outputs examined in the assessment—see Table 4.

The SBM model also offers an outline of the changes that really should be made to inputs and outputs to convert inefficient OPs into efficient ones—see Fig. 3.

The 'number of researchers working in improved R&I infrastructures' has the largest potential for improvement (2174%), followed by 'total eligible spending' (−78%), the 'number of enterprises supported for new-to-the-market products' (71%), the 'number of enterprises supported' (46%) and the 'number of enterprises working with R&I institutions'—see Fig. 3. All in all, like other studies, our findings also show the importance of the lack of skills as an obstacle to R&I OPs' implementation (e.g., Belitz & Lejpras, 2016; Duarte et al., 2017; Gardocka-Jałowiec & Wierzbicka, 2019). These results also highlight the need to foster the cooperation and networking of SMEs with research institutions, thus corroborating Hervás-Oliver et al. (2021) findings. Besides, additionally, since there seems to be an overuse of the EU funding (because of the required reduction on eligible spending), our results suggest the validation of the 'European paradox' since there seems to exist an 'innovation gap' in that supporting innovation inputs through public funding does not necessarily lead to innovation outputs (Hammadou et al., 2014; Radicic & Pugh, 2017).

Table 4 Characteristics of efficient OPs

DMU	Efficient score	Number of times as benchmark	Researchers working in improved infrastructures	Enterprises supported	Enterprises working with research institutions	Enterprises supported for new-to-the-market products	Total eligible spending	Innovation performance*
Alsace—ERDF	1.06	3	189.83	1909	46	24	19,104,006	Moderate innovator
Aragón—ERDF	1.47	17	340	696	267	0	1,399,896	Moderate innovator
Brussels Capital Region—ERDF	1.32	41	18,539	145	83	65	6,337,597	Innovation leader
Castilla y León—ERDF	1.11	7	333	0	3076	0	69,516,817	Moderate innovator
Competitiveness Entrepreneurship and Innovation—GR—ERDF/ESF	1.88	5	847	340	0	0	817,920	Moderate innovator
England—ERDF	1.63	5	78	19,146	4169	2475	547,408,267	Innovation leader
Extremadura—ERDF	1.23	6	3662	1181	1027	0	13,896,966	Emerging innovator
Multi-regional Spain—ERDF	1.25	0	255	562	159	0	42,966,355	Moderate innovator
Toscana—ERDF	1.16	18	238	4470	1216	2289	189,360,205	Strong innovator

(continued)

Table 4 (continued)

DMU	Efficient score	Number of times as benchmark	Researchers working in improved infrastructures	Enterprises supported	Enterprises working with research institutions	Enterprises supported for new-to-the-market products	Total eligible spending	Innovation performance*
Wallonia—ERDF	1.04	4	110	7232	1211	131	147,512,404	Strong innovator

* According to the Regional Innovation Scoreboard in 2021 (Hollanders, 2021)

Source Authors' own elaboration

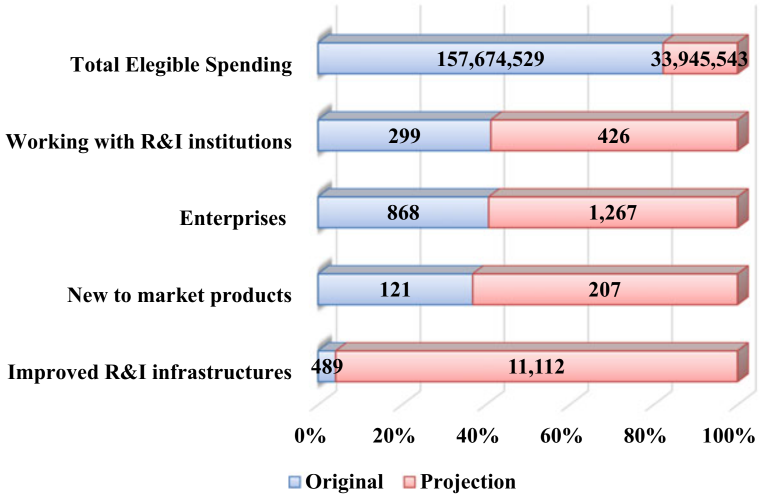


Fig. 3 Average original factors versus their projections for inefficient OPs. *Source* Authors' own elaboration

4.1 Results Obtained with SFA

To remove the potential effects of contextual factors and random errors, the functional forms given in (4) were estimated. The slacks of the outputs were considered as dependent variables, originating four regressions models. The multicollinearity was evaluated through the variance inflation factor (VIF), which measures the strength of correlation between the independent variables. This indicator is always greater than or equal to 1. Table 5 illustrates the values of VIF considering three sets of variables. When VIF is higher than 10, there is significant multicollinearity that needs to be corrected, thus, in the first step, we began by removing the four contextual variables that verify such condition. Afterward, the VIF values for the remaining variables were calculated (see Table 5, second step). Usually, values within 1 and 5 are not deemed relevant to cause concern (Belsley, 1991; James et al., 2013), but it seemed prudent to recalculate the VIF values without the variables "Non-R&D innovation expenditures" and "Innovation expenditures per person employed" and "PCT patent applications". The small values of VIF presented in Table 5 reveal that, when considering these variables, there is no problem of collinearity.

To run the SFA regression models, the R software, version 4.0.5 (RStudio Team, 2021), particularly, the sfaR package version 0.1.1 was used (Dakpo et al., 2022). The final regression models are shown in Table 6.

In model (1), the value of γ is very close to zero, thus the statistical noise is in a dominant position. Furthermore, statistical noise and the contextual variables $GDPPPP_{pc}$, Lifelong learning, and R&D expenditures of the public sector explain virtually all the variation that occurred in the slack of the output "Researchers working in improved infrastructures". For the models (2), (3), and (4), the values of γ are near

Table 5 VIF values

Contextual variables	VIF (1st step)	VIF (2nd step)	VIF (3rd step)
$GDPPPP_{pc}$	5.610	2.212	1.590
Population with tertiary education	2.611	1.879	1.678
Lifelong learning	3.955	3.290	1.962
R&D expenditures public sector	4.747	2.079	1.485
R&D expenditures business sector	6.034	3.497	1.604
Non-R&D innovation expenditures	4.489	4.050	–
Innovation expenditures per person employed	4.927	4.560	–
Product process innovators	11.337	–	–
Business process innovations	13.350	–	–
Innovative SMEs collaborating with others	8.155	3.769	2.317
PCT patent applications	7.489	4.203	–
Employment knowledge-intensive activities	12.199	–	–
Employment in innovative SMEs	24.199	–	–
Sales of new-to-market and new-to-firm innovations	2.848	2.617	1.691

Source Authors' own elaboration

one and statistically significant (1%), this means that management problems are the principal cause of the achieved technical (in)efficiency. The contextual variables considered in model (2) cause a relevant effect on the slack since all the regression coefficients associated are significant (at the 1% level). Likewise, in model (3), we found statistically significant variables to explain the required adjustments in "Enterprises Working with Research Institutions." Concerning model (4), since there are no statistically significant variables, no adjusted values are required for this output.

According to Table 6, a rise in GDPPPPpc contributes to a larger necessary increase of "Researchers Working in Improved Infrastructures", "Enterprises Supported" and "Enterprises Working with Research Institutions". On the one hand, regarding the two latter indicators, these findings seem to suggest that richer regions do not show a better use of ERDF targeted to strengthen R&I in SMEs. Bukvić et al. (2021) arrived at similar conclusions regarding the underuse of ERDF by SMEs in the Information and Communication Technologies sector in Croatia from 2014–2020. They ascertained that the difficulties and time required to submit, produce, and assess project proposals were a probable justification for these findings. Furthermore, Martinez-Cillero et al. (2020) reported that SMEs' investments are poorer than would be anticipated by standard economic models, proposing that these firms are particularly sensitive to funding difficulties. Another possible explanation might be related to the use of further financing opportunities in the framework of other funding programs (outside ERDF). On the other hand, regarding the first indicator,

Table 6 SFA analysis results

Variables	Slacks			
	Researchers working in improved infrastructures (1)	Enterprises supported (2)	Enterprises working with research institutions (3)	Enterprises supported for new-to-the-market products (4)
Constant	12,250.961	−34.988***	−87.114***	44.436
$GDPPPP_{pc}$	51.398*	0.389***	0.135**	−0.146
Population with tertiary education	–	127.327***	18.427***	−8.531
Lifelong learning	−16,132.525***	−108.351***	−36.908	8.531
R&D expenditures public sector	−8009.262*	64.496***	153.145***	−6.639
R&D expenditures business sector	3053.671	277.625***	−64.817*	24.805
Innovative SMEs collaborating with others	5076.931	−403.234***	−36.068	31.888
Sales of new-to-market and new-to-firm innovations	–	–	79.961***	−118.122
Sigma-squared	34,977,953***	2,289,881***	41,405***	120,379***
Gamma	0.003	0.99**	0.99***	0.99***
Log-likelihood function	−434.469	−346.055	−259.779	−282.7234

*, ** and *** Significance at the 10%, 5% and 1% levels, respectively
Source Authors' own elaboration

these outcomes also highlight the need to handle the lack of skilled researchers, a major hurdle to innovation also identified in more developed regions (Hölzl & Janger, 2014).

Additionally, a higher percentage of the population with tertiary education does not lead to an efficient number of "Enterprises Supported" and "Enterprises Working with Research Institutions" supported. These results might suggest that higher education institutions should be further contributing to the actual needs of the economy. In this framework, initiatives should be promoted to increase the relationship of SMEs with higher education institutions, since this type of linkage can be beneficial for the innovation environment (Kobarg et al., 2018; Rajalo & Vadi, 2017).

Lifelong learning seems to be positively associated with a better performance of the outputs "Researchers Working in Improved Infrastructures" and "Enterprises Supported" because it is negatively related to their required improvements. These outcomes may be explained by the abundance of the population involved in lifelong learning activities in the generality of MS (Anderson & Stejskal, 2019). Besides, these findings also suggest that coordinated lifelong learning policies play a pivotal role in propelling innovation and progress among MS and regions.

R&D expenditures within the public sector seem to have a positive contribution to the required enhancement of the adjustments on "Researchers Working in Improved Infrastructures", with the opposite effect on the number of "Enterprises Supported" and "Enterprises Working with Research Institutions". On the one hand, these results highlight the positive effect of public R&D spending on education attainment (i.e., a higher number of skilled researchers) since these are also linked with expenses in the higher education public sector. However, the two latter findings may imply that increased R&D expenditures within the public sector are not a viable strategy to mitigate SMEs' inability to engage in R&D (Hervás-Oliver et al., 2021). This might also suggest that EU SMEs cannot absorb the spillover effects from public R&D (Rodríguez-Pose & Wilkie, 2019). Furthermore, since the R&D expenditures within the business sector show a positive effect on the required enhancement of "Enterprises Working with Research Institutions" (i.e., a reduction of the necessary adjustment to become efficient), these findings appear to demonstrate the lesser influence of public R&D in SME innovation compared to private R&D. Similarly, results were also attained by Hervás-Oliver et al. (2021).

In what concerns the innovative SMEs collaborating with others as a percentage of SMEs, there is a positive effect both on the number of "Enterprises Supported" and the "Enterprises Working with Research Institutions" (the enhancement required in these two outputs is negative). In a similar context, Hervás-Oliver et al. (2021) concluded that SME collaboration with exterior sources of knowledge (either supply-chain actors and competitors or universities or other sources of research) is positively related to regional SME innovation.

Finally, sales of new-to-market and new-to-firm innovations require a further enhancement of the number of "Enterprises Working with Research Institutions" (the enhancement required in this output is positive). These findings might be influenced by the fact that this contextual variable does not make a distinction between incremental and radical innovation, also considering non-technological innovations (Apa et al., 2021).

4.2 Results Obtained with the Adjusted Factors

Table 7 shows that efficient OPs hardly change their average efficiency scores with the adjusted factors (the standard deviation is the same, i.e., 0.27). Besides, the efficiency scores are bounded within the same interval, i.e., [1.04, 1.88], demonstrating efficiency scores bigger than 1.24 for more than 50% of the efficient OPs. Also, inefficient OPs decrease the variability of their efficiency scores (with a standard deviation of 0.11 against the previous 0.15, with more than 50% of inefficient OPs having efficiency values just under 0.06) and increased their average efficiency from 0.06 to 0.10 (underlining the importance of the contextual variables).

Table 7 Descriptive statistics of the results obtained for efficient and inefficient OPs with adjusted factors

	Statistics	Efficiency score	Researchers working in improved infrastructures	Enterprises supported	Enterprises working with research institutions	Enterprises supported for new to market products	Total eligible spending
Efficient OPs	Mean	1.31	3,917.39	5,436.94	1,109.54	498.40	227,352,595.60
	Median	1.24	336.50	1,545.00	647.00	12.00	44,310,411.50
	Standard deviation	0.27	6,870.73	7,597.35	1,438.08	994.60	405,454,176.09
	Minimum	1.04	77.68	0.00	0.00	0.00	817,920.00
	Maximum	1.88	18,538.90	19,250.00	4,169.00	2,475.00	1,278,171,878.00
	Count	10	10	10	10	10	10
Inefficient OPs	Mean	0.10	1,685.23	945.29	356.66	120.84	157,674,529.09
	Median	0.06	2,092.75	380.50	165.93	15.00	50,000,231.00
	Standard deviation	0.11	984.28	1,784.74	638.86	197.88	405,563,908.88
	Minimum	0.00	10.00	0.00	0.00	0.00	1,424,378.00
	Maximum	0.43	2,840.20	9,677.00	3,590.18	660.00	2,654,718,365.00
	Count	43	43	43	43	43	43

Source Authors' own elaboration

Figure 4 depicts the difference in the technical efficiency of the OPs with and without adjusted factors.

Figure 5 illustrates the greatest efficiency gains attained with the adjusted factors. When contrasted with the first step of the analysis, “Competitiveness and Cohesion–HR—ERDF/CF” demonstrated the greatest gain in efficiency, with values going from 0.0003 to 0.0676. Overall, these outcomes suggest that the inefficiencies originally computed for these OPs were not solely the result of their low technical level but were also related to their contextual factors.

Then again, out of the ten efficient OPs, the three most chosen as benchmarks are “Brussels Capital Region—ERDF” (39 times), “Aragón—ERDF” (20 times), and “Toscana—ERDF” (17 times) and “Extremadura—ERDF” (15 times).

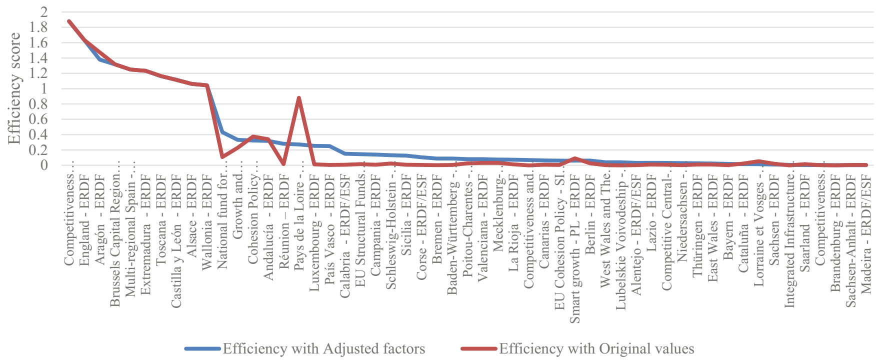


Fig. 4 Efficiency scores for the efficient OPs obtained with adjusted and non-adjusted factors. *Source* Authors’ own elaboration

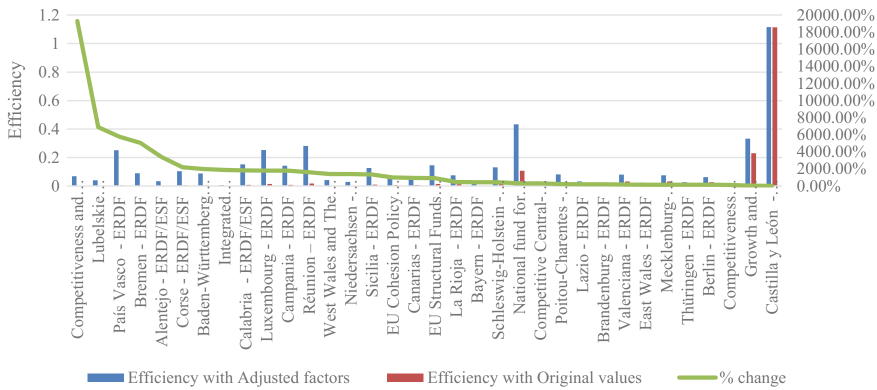


Fig. 5 Efficiency scores for the efficient OPs with the greatest efficiency gains obtained with adjusted and non-adjusted factors. *Source* Authors’ own elaboration

5 Conclusions and Further Research

The primary purpose of this article was to evaluate the reasons behind the inefficiency of the OPs devoted to boosting R&I in SMEs. With this aim, we assessed 53 OPs within TO1 from 19 EU MS. To begin with, the SBM modeling approach is utilized to calculate the technical efficiency of every OP. At this stage, important data about the overall adjustments that should be made to reduce any disparities between inefficient OPs and their corresponding benchmarks are obtained.

Unlike other commonly used techniques applied in comparable situations, such as reference cases, econometric and statistical methods, and macroeconomic and microeconomic analyses, the SBM model can be particularly useful for MA, as it enables them to identify the references of best practices and the required changes to improve the OPs' implementation performance, also contemplating their performance in two different stages. The second phase consists of employing SFA to the slacks of inefficient OPs to change the inputs and outputs after removing environmental effects and statistical noise. At this stage, information is extracted about how environmental factors may influence the efficiency of ERDF deployment in distinct OPs devoted to the promotion of R&I in SMEs, as well as the magnitude of management flaws. Finally, the previously corrected factors are employed in the SBM model to obtain new efficiency ratings.

Our main conclusions are discussed next.

RQ1: "Which contextual variables show a relevant effect on the inefficiencies of the OPs committed to boosting R&I in SMEs?"

Our results indicate that more developed regions do not make efficient use of ERDF aimed at promoting R&I in SMEs. The difficulty and time necessary to submit, develop and evaluate the project proposals, and the higher vulnerability of these types of enterprises to financial issues are possible explanations for these poor results. Alternatively, these findings can also be attributed to the utilization of additional financing options within the context of other funding programs. Furthermore, these results also demonstrate the need of addressing the shortage of trained researchers, which has been recognized as a key barrier to innovation in more developed regions. Besides, a larger percentage of the population with university education does not result in an adequate number of "Enterprises" and "Enterprises Working with Research Institutions" supported. These findings may imply that higher education institutions should contribute more to the economy's genuine demands. Initiatives should be pushed in this framework to strengthen SMEs' relationships with higher education institutions since this form of collaboration can be advantageous to the innovation environment.

Lifelong learning appears to be favorably correlated with the higher performance of the outputs "Researchers Working in Improved Infrastructures" and "Enterprises Supported". Hence, our results indicate that integrated lifelong learning strategies are critical in accelerating innovation among MS and regions.

R&D expenditures in the public sector appear to contribute positively to the needed enhancement of the adjustments on "Researchers Working in Improved Infrastructures" but have the opposite effect on the number of "Enterprises Supported" and "Enterprises Working with Research Institutions". On the one hand, these findings indicate the favorable impact of these expenditures on educational attainment because they are also connected to expenditures in the public higher education sector. The two latter findings, however, may suggest that greater R&D expenditures in the public sector are not a realistic option for mitigating SMEs' incapacity to engage in R&D. This might also imply that SMEs cannot absorb spillover effects from governmental R&D.

Additionally, because R&D expenditures in the business sector have a beneficial impact on the required improvement of "Enterprises Working with Research Institutions", these findings appear to show that public R&D has a lesser influence on SME innovation compared to private R&D.

Concerning innovative SMEs cooperating with others as a proportion of SMEs, there is a favorable influence of this contextual variable both on the number of "Enterprises" and on "Enterprises Working with Research Institutions" supported. Therefore, it might be ascertained that the collaboration of SMEs with external sources of knowledge is positively connected to regional SME innovation.

Finally, sales of new-to-market and new-to-firm innovations have a negative effect on the number of "Enterprises Working with Research Institutions" (the enhancement required in this output is positive). These findings might be impacted by the fact that this contextual variable does not distinguish between incremental and radical innovation, as well as non-technological breakthroughs.

RQ2: "What are the impacts of considering contextual factors on the efficiency of the OPs?"

If the factors are adjusted according to Tone and Tsutsui (2009), 19% of OPs (10) manage to attain technical efficiency in any case, indicating that the effects of contextual factors are more visible on inefficient OPs. The biggest gap in efficiency was found in inefficient OPs that showed an average efficiency gain of 67%. The most efficient regions regardless of the adjustments were "Competitiveness Entrepreneurship and Innovation—GR—ERDF/ESF", "England—ERDF" and "Multi-regional Spain—ERDF", with values of efficiency ranging between 1.25 and 1.88. In general, it can be concluded that the technical efficiency of the OPs classified as efficient was mostly driven by good management practices.

Although this work gave new perspectives on the evaluation of OPs dedicated to R&I in EU SMEs, it had limitations. Though the performance framework made available by the European Commission includes a set of procedural indicators, there is no full correspondence between the data collected for the OPs' accomplishments and their financial execution. Secondly, since the data available is often sparse, our evaluation was applied to a reduced number of OPs.

Whereas our work focused on an assessment technique for use throughout the reporting phases of the programmatic cycle, future work should address ex-post assessment, with a special emphasis on the spillover effects of the OPs under TO1.

Acknowledgements This work has been funded by European Regional Development Fund in the framework of Portugal 2020—Programa Operacional Assistência Técnica (POAT 2020), under project POAT-01-6177-FEDER-000044 ADEPT: Avaliação de Políticas de Intervenção Co-financiadas em Empresas. INESC Coimbra and CeBER are supported by the Portuguese Foundation for Science and Technology funds through Projects UID/MULTI/00308/2020 and UIDB/05037/2020, respectively.

References

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics, 6*(1), 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

Anderson, H. J., & Stejskal, J. (2019). Diffusion efficiency of innovation among EU member states: A data envelopment analysis. *Economies, 7*(2), 34. https://doi.org/10.3390/economies7020034

Apa, R., De Marchi, V., Grandinetti, R., & Sedita, S. R. (2021). University-SME collaboration and innovation performance: The role of informal relationships and absorptive capacity. *The Journal of Technology Transfer, 46*(4), 961–988. https://doi.org/10.1007/s10961-020-09802-9

Athanassopoulos, A. D. (1996). Assessing the comparative spatial disadvantage (CSD) of regions in the European Union using non-radial data envelopment analysis methods. *European Journal of Operational Research, 94*(3), 439–452. https://doi.org/10.1016/0377-2217(95)00114-X

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Barbero, J., Diukanova, O., Gianelle, C., Salotti, S., & Santoalha, A. (2022). Economic modelling to evaluate smart specialisation: An analysis of research and innovation targets in Southern Europe. *Regional Studies*, 1–14. https://doi.org/10.1080/00343404.2021.1926959

Bedu, N., & Vanderstocken, A. (2020). Do regional R&D subsidies foster innovative SMEs' development: Evidence from Aquitaine SMEs. *European Planning Studies, 28*(8), 1575–1598. https://doi.org/10.1080/09654313.2019.1651828

Belitz, H., & Lejpras, A. (2016). Financing patterns of R&D in small and medium-sized enterprises and the perception of innovation barriers in Germany. *Science and Public Policy, 43*(2), 245–261. https://doi.org/10.1093/scipol/scv027

Belsley, D. A. (1991). *Conditioning diagnostics: Collinearity and weak data in regression* (p. 1991). Wiley.

Berkowitz, P., Monfort, P., & Pieńkowski, J. (2019). Unpacking the growth impacts of European Union Cohesion Policy: Transmission channels from Cohesion Policy into economic growth. *Regional Studies*. https://doi.org/10.1080/00343404.2019.1570491

Bianchi, M., Campodall'Orto, S., Frattini, F., & Vercesi, P. (2010). Enabling open innovation in small-and medium-sized enterprises: How to find alternative applications for your technologies. *R&D Management, 40*(4), 414–431. https://doi.org/10.1111/j.1467-9310.2010.00613.x

Bouncken, R. B., Fredrich, V., & Kraus, S. (2020). Configurations of firm-level value capture in coopetition. *Long Range Planning, 53*(1), 101869. https://doi.org/10.1016/j.lrp.2019.02.002

Bukvić, I. B., Babić, I. Đ, & Starčević, D. P. (2021). Study on the Utilization of National and EU Funds in Financing Capital Investments of ICT Companies. In 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 1282–1287). IEEE. https://doi.org/10.23919/MIPRO52101.2021.9597077.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research, 2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Dahlander, L., & Gann, D. M. (2010). How open is innovation? *Research Policy, 39*(6), 699–709. https://doi.org/10.1016/j.respol.2010.01.013

Dakpo, K. H., Desjeux, Y., & Latruffe L. (2022). sfaR: Stochastic Frontier Analysis using R. *R Package Version 0.1.1.*

Di Comite, F., Lecca, P., Monfort, P., Persyn, D., & Piculescu, V. (2018). The impact of Cohesion Policy 2007–2015 in EU regions: Simulations with the RHOMOLO Interregional Dynamic General Equilibrium Model (No. 03/2018). JRC Working Papers on Territorial Modelling and Analysis. Retrieved April 11, 2022, from https://www.econstor.eu/bitstream/10419/202268/1/jrc-wptma201803.pdf

Diukanova, O., Mandras, G., & Di Comite, F. (2022). Modelling the effects of R&I and low-carbon european structural funds: The case of Apulia, Italy. *Scienze Regionali, 21*(Speciale), 9–38. https://doi.org/10.14650/103213

Duarte, F. A., Madeira, M. J., Moura, D. C., Carvalho, J., & Moreira, J. R. M. (2017). Barriers to innovation activities as determinants of ongoing activities or abandoned. *International Journal of Innovation Science, 9*, 244–264. https://doi.org/10.1108/IJIS-01-2017-0006

Durlauf, S. N. (2009). The rise and fall of cross-country growth regressions. *History of Political Economy, 41*(Suppl_1), 315–333. https://doi.org/10.1215/00182702-2009-030

Fattorini, L., Ghodsi, M., & Rungi, A. (2020). Cohesion policy meets heterogeneous firms. *JCMS: Journal of Common Market Studies, 58*(4), 803–817. https://doi.org/10.1111/jcms.12989

Fried, H. O., Lovell, C. K., Schmidt, S. S., & Yaisawarng, S. (2002). Accounting for environmental effects and statistical noise in data envelopment analysis. *Journal of Productivity Analysis, 17*(1), 157–174. https://doi.org/10.1023/A:1013548723393

García-Quevedo, J., Segarra-Blasco, A., & Teruel, M. (2018). Financial constraints and the failure of innovation projects. *Technological Forecasting and Social Change, 127*, 127–140. https://doi.org/10.1016/j.techfore.2017.05.029

Gardocka-Jałowiec, A., & Wierzbicka, K. (2019). Barriers to creating innovation in the polish economy in the years 2012–2016. *Studies in Logic, Grammar and Rhetoric, 59*(1). https://doi.org/10.2478/slgr-2019-0038

Gómez-García, J., Enguix, M. D. R. M., & Gómez-Gallego, J. C. (2012). Estimation of the efficiency of structural funds: A parametric and nonparametric approach. *Applied Economics, 44*(30), 3935–3954. https://doi.org/10.1080/00036846.2011.583224

Gouveia, M. C., Henriques, C. O., & Costa, P. (2021). Evaluating the efficiency of structural funds: An application in the competitiveness of SMEs across different EU beneficiary regions. *Omega, 101*, 102265. https://doi.org/10.1016/j.omega.2020.102265

Gramillano, A., Celotti, P., Familiari, G., Schuh, B., & Nordstrom, M. (2018). Development of a system of common indicators for European Regional Development Fund and Cohesion Fund interventions after 2020. Study for the DG for Regional and Urban Policy, European Commission. *Gen. Reg. Urban Policy Eur. Comm., 10*, 279688. Retrieved April 11, 2022, from https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/indic_post2020/indic_post2020_p2_en.pdf

Gustafsson, A., Tingvall, P. G., & Halvarsson, D. (2020). Subsidy entrepreneurs: An inquiry into firms seeking public grants. *Journal of Industry, Competition and Trade, 20*(3), 439–478. https://doi.org/10.1007/s10842-019-00317-0

Hammadou, H., Paty, S., & Savona, M. (2014). Strategic interactions in public R&D across European countries: A spatial econometric analysis. *Research Policy, 43*(7), 1217–1226. https://doi.org/10.1016/j.respol.2014.01.011

Henriques, C., Viseu, C., Neves, M., Amaro, A., Gouveia, M., & Trigo, A. (2022b). How efficiently does the EU support research and innovation in SMEs? *Journal of Open Innovation: Technology, Market, and Complexity, 8*(2), 92. https://doi.org/10.3390/joitmc8020092

Henriques, C., Viseu, C., Trigo, A., Gouveia, M., & Amaro, A. (202a). How efficient is the cohesion policy in supporting small and mid-sized enterprises in the transition to a low-carbon economy? *Sustainability, 14*(9), 5317. https://doi.org/10.3390/su14095317

Hervás-Oliver, J. L., Parrilli, M. D., Rodríguez-Pose, A., & Sempere-Ripoll, F. (2021). The drivers of SME innovation in the regions of the EU. *Research Policy, 50*(9), 104316. https://doi.org/10.1016/j.respol.2021.104316

Hollanders, H. (2021). *Regional Innovation Scoreboard 2021*. European Commission. https://doi.org/10.2873/674111

Hölzl, W., & Janger, J. (2014). Distance to the frontier and the perception of innovation barriers across European countries. *Research Policy, 43*(4), 707–725. https://doi.org/10.1016/j.respol.2013.10.001

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: Springer. https://doi.org/10.1007/978-1-4614-7138-7

Jondrow, J., Lovell, C. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics, 19*(2–3), 233–238. https://doi.org/10.1016/0304-4076(82)90004-5

Kobarg, S., Stumpf-Wollersheim, J., & Welpe, I. M. (2018). University-industry collaborations and product innovation performance: The moderating effects of absorptive capacity and innovation competencies. *The Journal of Technology Transfer, 43*(6), 1696–1724. https://doi.org/10.1007/s10961-017-9583-y

Lee, S., Park, G., Yoon, B., & Park, J. (2010). Open innovation in SMEs—An intermediated network model. *Research Policy, 39*(2), 290–300. https://doi.org/10.1016/j.respol.2009.12.009

Lopez-Rodríguez, J., & Faíña, A. (2014). Rhomolo and other methodologies to assess The European Cohesion Policy. *Investigaciones Regionales-Journal of Regional Research,* (29), 5–13. Retrieved April 11, 2022, from https://www.redalyc.org/pdf/289/28932224001.pdf

Martinez-Cillero, M., Lawless, M., O'Toole, C., & Slaymaker, R. (2020). Financial frictions and the SME investment gap: New survey evidence for Ireland. *Venture Capital, 22*(3), 239–259. https://doi.org/10.1080/13691066.2020.1771826

Marzinotto, B. (2012). The growth effects of EU cohesion policy: A meta-analysis (No. 2012/14). Bruegel Working Paper. Retrieved April 12, 2022, from https://www.econstor.eu/bitstream/10419/78011/1/728570688.pdf

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 435–444. https://doi.org/10.2307/2525757

Müller, J. M., Buliga, O., & Voigt, K. I. (2021). The role of absorptive capacity and innovation strategy in the design of industry 4.0 business Models-A comparison between SMEs and large enterprises. *European Management Journal, 39*(3), 333–343. https://doi.org/10.1016/j.emj.2020.01.002

Neto, P., & Santos, A. (2020). Guidelines for territorial impact assessment applied to regional research and innovation strategies for smart specialisation. In *Territorial Impact Assessment* (pp. 211–230). Springer. https://doi.org/10.1007/978-3-030-54502-4_12

Ortiz, R., & Fernandez, V. (2022). Business perception of obstacles to innovate: Evidence from Chile with pseudo-panel data analysis. *Research in International Business and Finance, 59*, 101563. https://doi.org/10.1016/j.ribaf.2021.101563

R Core Team. (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. Retrieved from https://www.R-project.org

Radicic, D., & Pugh, G. (2017). R&D programmes, policy mix, and the 'European paradox': Evidence from European SMEs. *Science and Public Policy, 44*(4), 497–512. https://doi.org/10.1093/scipol/scw077

Rajalo, S., & Vadi, M. (2017). University-industry innovation collaboration: Reconceptualization. *Technovation, 62–63*, 42–54. https://doi.org/10.1016/j.technovation.2017.04.003

Rodríguez-Pose, A., & Wilkie, C. (2019). Innovating in less developed regions: What drives patenting in the lagging regions of Europe and North America. *Growth and Change, 50*(1), 4–37. https://doi.org/10.1111/grow.12280

Romero-Martínez, A. M., Fernández-Rodríguez, Z., & Vázquez-Inchausti, E. (2010). Exploring corporate entrepreneurship in privatized firms. *Journal of World Business, 45*(1), 2–8. https://doi.org/10.1016/j.jwb.2009.04.008

Santos, A., Cincera, M., Neto, P., & Serrano, M. M. (2019). Which projects are selected for an innovation subsidy? The Portuguese case. *Portuguese Economic Journal, 18*(3), 165–202. https://doi.org/10.1007/s10258-019-00159-y

Sein, Y. Y., & Prokop, V. (2021). Mediating role of firm R&D in creating product and process innovation: Empirical evidence from Norway. *Economies, 9*(2), 56. https://doi.org/10.3390/economies9020056

Stojčić, N., Srhoj, S., & Coad, A. (2020). Innovation procurement as capability-building: Evaluating innovation policies in eight Central and Eastern European countries. *European Economic Review, 121*, 103330. https://doi.org/10.1016/j.euroecorev.2019.103330

Thum-Thysen, A., Voigt, P., Bilbao-Osorio, B., Maier, C., & Ognyanova, D. (2019). Investment dynamics in Europe: Distinct drivers and barriers for investing in intangible versus tangible assets? *Structural Change and Economic Dynamics, 51*, 77–88. https://doi.org/10.1016/j.strueco.2019.06.010

Tone, K., & Tsutsui, M. (2009). Tuning regression results for use in multi-stage data adjustment approach of DEA (<Special Issue> Operations Research for Performance Evaluation). *Journal of the Operations Research Society of Japan, 52*(2), 76–85. https://doi.org/10.15807/jorsj.52.76

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research, 130*(3), 498–509. https://doi.org/10.1016/S0377-2217(99)00407-5

Tone, K. (2002). A slacks-based measure of super-efficiency in data envelopment analysis. *European Journal of Operational Research, 143*(1), 32–41. https://doi.org/10.1016/S0377-2217(01)00324-1

Van de Vrande, V., De Jong, J. P., Vanhaverbeke, W., & De Rochemont, M. (2009). Open innovation in SMEs: Trends, motives and management challenges. *Technovation, 29*(6–7), 423–437. https://doi.org/10.1016/j.technovation.2008.10.001

Wostner, P., & Šlander, S. (2009). The effectiveness of EU cohesion policy revisited: are EU funds really additional? University of Strathclyde. Retrieved April 12, 2022, from https://strathprints.strath.ac.uk/70313/1/EPRP_69.pdf

Zhou, X., Rasool, S. F., Yang, J., & Asghar, M. Z. (2021). Exploring the relationship between despotic leadership and job satisfaction: The role of self-efficacy and leader–member exchange. *International Journal of Environmental Research and Public Health, 18*(10), 5307. https://doi.org/10.3390/ijerph18105307

Zimmermann, V., & Thomä, J. (2016). SMEs face a wide range of barriers to innovation-support policy needs to be broad-based. *KfW Research Focus on Economics, 130*, 1–8.

Are ERDF Devoted to Boosting ICT in SMEs Inefficient? Insights Through Different DEA Models

Carla Henriques and Clara Viseu

Abstract We evaluated the execution of operational programs (OPs) targeted at increasing the adoption of information and communication technologies (ICT) in small and midsize enterprises (SMEs). As a result, we employed two different data envelopment analysis (DEA) models to evaluate 51 OPs from 16 countries, contemplating data provided to the European Union (EU). All in all, we observed that almost 20% of the OPs (10) achieved efficient procedural outcomes, with the slack-based method (SBM) and with the weighted Russel Directional Distance model (WRDDM), respectively. Two of the OPs most frequently viewed as benchmarks were in Spain (the country that uses "vouchers" simplifying processes), remaining robustly efficient for data perturbations of 5% and 10%. The 'number of operations supported' is the metric that necessitates more consideration according to both models. Overall, these findings show the robustness of results with both methods, highlighting a higher discriminatory power of the second method, particularly for inefficient OPs. Finally, the unsuccessful findings attained might be linked to bureaucratic procedures and SMEs' incapacity to deal with the complicated processes involved in obtaining and implementing European Regional Development Fund (ERDF) proposals. As a result, it is critical to provide the extra help that reduces managerial requirements while also meeting the demands of SMEs.

Keywords ICT · SMEs · SBM model · WRDD model · ERDF

C. Henriques (✉) · C. Viseu
Polytechnic of Coimbra, Coimbra Business School, 3045-601 Coimbra, Portugal
e-mail: chenriques@iscac.pt

C. Henriques
INESC Coimbra, Department of Electrical and Computer Engineering (DEEC), University of Coimbra, Polo 2, 3030-290 Coimbra, Portugal

Centre for Business and Economics Research (CeBER), Faculty of Economics, University of Coimbra, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

C. Viseu
Centro de Matemática e Aplicações (CMA-UBI), Universidade da Beira Interior, 6200-001 Covilhã, Portugal

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_2

1 Introduction

Various efforts have been undertaken over the last ten years to investigate the significance of ICT use in SMEs, leading to a huge array of published studies. Taylor (2015) explored two theoretical foundations in this respect: ‘diffusion of innovation theory’ (Rogers et al., 2014) and ‘technology, organization, and environment structure’ (Tornatzky and Fleischer, 1990) to provide a thorough solid background for SMEs’ usage of ICT. This comprehensive framework contains a taxonomy that covers several of the most important internal and external variables influencing SMEs’ ICT adoption. Mbuyisa and Leonard (2017), on the other hand, studied the link between ICT, SMEs, and poverty alleviation. To begin, this study emphasizes the significance of ICT deployment in SMEs. Following that, it investigates how SMEs might use ICTs to help them eliminate poverty. Disparities in ICT availability among organizations were also investigated in a study on the digital divide (DD) among firms conducted by Bach et al (2013). This assessment looked at the region, firm type, and research period, and the impact and causes of DD. Another study examined the key reasons, repercussions, and barriers to ICT adoption in SMEs (Consoli, 2012). Tarutė and Gatautis (2014) evaluated the possible impact of ICT on SMEs’ competitiveness. Lehner and Sundby (2018) examined the importance of ICT skills and competencies for SMEs from several perspectives.

Oberländer et al. (2020) assessed the subject of workplace digital competencies. Isensee et al. (2020) developed a conceptual model to emphasize the connections among three domains, to present the first comprehensive picture of organizational behavior, sustainability, and digitalization levels in SMEs, as well as their interconnections. They found that their strategies, organizational key competencies, governance, and mindsets are by far the most explored cultural features. Another recent study (Ramdani et al, 2022) looked at digital innovation in SMEs and discovered that it is driven by past experiences, progresses throughout many stages of innovation, and results in organizational and corporate ongoing learning. Similarly, Thrassou et al. (2020) did a literature review to help identify the major problems and possibilities for SMEs in the framework of digitization and ICT advancements.

Overall, none of the previous studies focuses on the evaluation of ICT policy. According to Reggi & Gil-Garcia (2021), there has been a lack of scholarly attention on how to select the best approaches for allocating funding amongst diverse ICT initiatives. Moreover, limited researchers have attempted to determine if this sort of funding is delivered in line with some of the most pressing requirements of each place (Kleibrink et al., 2015; Reggi & Scicchitano, 2014). Ex-post assessments are frequently utilized in research evaluating European structural funding awarded to ICT (Kleibrink et al., 2015; Reggi & Scicchitano, 2014). There is additional research that provides an ex-ante assessment of the factors influencing funding distribution among various ICT activities (Reggi & Gil-Garcia, 2021).

Nonetheless, no research has been undertaken so far that compares the implementation of OPs connected to ICT policies to their peers throughout the same period, or that indicates the modifications that must be applied to achieve the OPs’ efficiency.

Consequently, the objective of this paper is to contribute to the existing literature by employing a mathematical method that allows Management Authorities (MA) to better assess the implementation of OPs focused on ICT deployment in SMEs by using two non-parametric methodologies, specifically the SBM and the WRDD models. In this context, the research questions (RQ) that this work seeks to answer are as follows:

RQ1: "What indicators restrict the successful use of ERDF allocated to enhance ICT adoption in EU SMEs?"

RQ2: "What OPs are generally elected as benchmarks across the assessed time frame?"

RQ3: "What OPs demonstrate more resilience in terms of their efficiency classification in the context of changes in the indicators used?"

The following is the structure for this paper. Section 2 describes the fundamental principles behind the methodologies proposed to evaluate the implementation of the OPs under consideration. Section 3 discusses the main reasons for selecting the inputs and outputs, also presenting their descriptive statistics. The main results are discussed in Sect. 4. Section 5 summarizes the important findings, examines prospective political implications, identifies critical weaknesses, and suggests future research directions.

2 Methodology

Classical DEA techniques, such as the CCR (Charnes et al., 1978) and BCC (Banker et al., 1984), are radial, which means that they can only manage proportional changes in the inputs (resources) or outputs (outcomes). Consequently, the CCR and BCC efficiency ratings produced indicate the highest proportionate input (output) contraction (expansion) rates for all inputs (outputs). Unfortunately, this sort of premise is frequently not met. As a result, we utilize the SBM model (Tone, 2001), which provides for a much more extensive evaluation of efficiency since it is non-radial (i.e., inputs and outputs can adapt non-proportionally) and non-oriented (i.e., enables addressing simultaneous changes of the inputs and outputs). In addition, we applied the WRDDM (Chen et al., 2015), which allows us to decompose inefficiencies.

2.1 The SBM and the WRDD Models

Consider a set of n DMUs ($\mathrm{DMU}_1, \mathrm{DMU}_2, \ldots, \mathrm{DMU}_\mathrm{n}$) with $X = [x_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n]$ the $(m \times n)$ matrix of *inputs*, $Y = [y_{rj}, r = 1, 2, \ldots, s, j = 1, 2, \ldots, n]$ the matrix of *outputs* $(s \times n)$ and the rows of these matrices for DMU_k are, respectively, $\boldsymbol{x}_k^T$ and $\boldsymbol{y}_k^T$,, where $^\mathrm{T}$ is the transpose of a vector. Also, assume a Variable Returns to Scale (VRS) technology with the imposition of $\sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \geq 0(\forall_j)$..

The SBM model

$$\underset{\lambda, \boldsymbol{s}^-, \boldsymbol{s}^+}{Min} \rho = \frac{1-\frac{1}{m}\sum_{i=1}^{m} s_i^-/x_{ik}}{1+\frac{1}{s}\sum_{r=1}^{s} s_r^+/y_{rk}} \quad (1)$$

s.t.

$$x_{ik} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, \ i = 1, \ldots, m$$

$$y_{rk} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, \ r = 1, \ldots, s$$

$$\sum_{j=1}^{n} \lambda_j = 1, \ \lambda_j \geq 0, \ j = 1, \ldots, n$$

$$\lambda_j \geq 0, \ j = 1, \ldots, n$$

$$s_i^- \geq 0, i = 1, \ldots, m$$

$$s_i^+ \geq 0, r = 1, \ldots, s$$

Where the value of $0 < \rho < 1$ can be seen as the ratio of average inefficiencies of inputs and outputs. A DMU$_k$ is SBM-efficient if $\rho^* = 1$, meaning that the slacks (s_i^- and s_i^+) are null for all the inputs and outputs

The WRDDM

$$\max\beta_k^R = max\left(w_y\left(\sum_r \varpi_y^r\alpha_k^r\right) + w_x\left(\sum_i \varpi_x^i\zeta_k^i\right)\right) \quad (2)$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{rk} + \alpha_k^r g_{yr}, r = 1, \ldots, s,$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{ik} - \zeta_k^i g_{xi}, \ i = 1, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j = 1, \ \lambda_j \geq 0, \ j = 1, \ldots, n,$$

Where α_o^r and ζ_o^i are the inefficiency values for every output and input, respectively, and β_k^R is the overall inefficiency. $\zeta_o^i g_{xi}$ is the reduction needed on x_{ik} and $\alpha_o^r g_{yr}$ is the needed increase on y_{rk} to turn DMU$_k$ into an efficient DMU. The directional vectors $\boldsymbol{g}_x$ and $\boldsymbol{g}_y$ are such that $\left(-\boldsymbol{g}_x, \boldsymbol{g}_y\right) = \left(-\boldsymbol{x}_k, \boldsymbol{y}_k\right)$ A DMU$_k$ is efficient when $\beta_k^R = 0$. w_y and w_x, with $w_y + w_x = 1$, are weights that assign the importance of the outputs and inputs. The importance of the inefficiencies of every factor is defined such that $\sum_{r\epsilon O} \varpi_y^r = 1, \sum_{i\epsilon I} \varpi_x^i = 1$

Model (1) can be transformed into a linear problem, by using a positive scalar variable t (for further details see Tone (2001)).

The set of benchmark DMUs for each inefficient DMU$_k$ in both models is $E_k = \left\{j : \lambda_j^* > 0, j = 1, \ldots, n\right\}$ and the reference point for each inefficient DMU$_k$ is obtained as:

$$\left(\hat{\boldsymbol{x}}_k, \hat{\boldsymbol{y}}_k\right) = \left(\boldsymbol{x}_k - \boldsymbol{s}^{-*}, \boldsymbol{y}_k + \boldsymbol{s}^{+*}\right) = \left(\sum_{j\epsilon E_k} \lambda_j^*\boldsymbol{x}_j, \sum_{j\epsilon E_k} \lambda_j^*\boldsymbol{y}_j\right). \quad (3)$$

The WRDDM model's measure of inefficiency may be translated into a slack based measure departing from problem (4):

$$
\begin{aligned}
&\max\left(w_y\left(\sum_r \varpi_y^r \frac{s_r^{+'}}{g_{yr}}\right) + w_x\left(\sum_i \varpi_x^i \frac{s_i^{-'}}{g_{xi}}\right)\right)\\
&\text{s.t.} \sum_{j=1}^{n} \lambda_j y_{rj} = y_{ro} + s_r^{+'}, r = 1, \ldots, s,\\
&\sum_{j=1}^{n} \lambda_j x_{ij} = x_{io} - s_i^{-'}, i = 1, \ldots, m,\\
&\sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \geq 0,\ j = 1, \ldots, n,\\
&s_r^{+'} \geq 0(\forall_r), s_i^{-'} \geq 0(\forall_i).
\end{aligned}
\tag{4}
$$

Let $\left(s_r^{+*'},\ s_i^{-*'},\ \lambda_j^*\right)$ be the optimal solution to problem (4). The WRDDM inefficiency measure can be decomposed from:

$$
\left(w_y\left(\sum_r \alpha_o^{r*'}\right) + w_x\left(\sum_i \zeta_o^{i*'}\right)\right), \text{ where } \alpha_o^{r*'} = \varpi_y^r \frac{s_r^{+'}}{g_{yr}} \cdot \text{e} \cdot \zeta_o^{i*'} = \varpi_x^i \frac{s_i^{-'}}{g_{xi}} \tag{5}
$$

3 Data

The input and output indicators used for measuring the efficiency of ERDF dedicated to ICT adoption in SMEs were picked from a set of measures mandated by the EU.[1]

3.1 Input and Output Factors

3.1.1 "Total Eligible Costs Decided" and "Total Eligible Spending"

The criteria used to measure the capacity of the OPs' absorption are "total eligible spending" and "total eligible costs decided". The former concerns qualified costs that have been documented and verified by a decision authority. Therefore, this component is employed as an output since the more the value assigned to it, the greater the financial implementation of each operation. The latter is regarded as an input since it relates to the resources that are available to the initiatives selected for support, which must be maintained as low as possible.

[1] Available online: https://cohesiondata.ec.europa.eu/2014-2020-Categorisation/ESIF-2014-2020-categorisation-ERDF-ESF-CF-planned-/3kkx-ekfq (accessed 30th March 2022).

3.1.2 Number of Operations Supported

The “number of operations supported” alludes to the number of projects financed by the ERDF. The greater the number of projects funded, the greater the prospect of improving organizational ICT usage. Consequently, this indicator is an output.

Table 1 shows the descriptive statistics of the inputs and outputs used in the OPs’ efficiency appraisal.

4 Discussion of Results

Results for the SBM model were obtained using Max DEA software, whereas for the WRDD model were obtained through an Excel Visual Basic-based application created by the authors for solving our DEA models that uses Excel Solver as the backend. The descriptive statistics thus obtained with both models are shown in Fig. 1.

Figure 1 shows, as expected, that efficient OPs have larger mean scores than inefficient ones. Nevertheless, there is a lower variability of the scores obtained for efficient OPs with the WRDD model when compared to those computed with the SBM model (efficiency within the range [1, 1.31] and [1.00, 1.71] and at least 50% of efficient OPs having efficiency values greater 1.07 and 1.16 according to the WRDD and SBM models, respectively). Besides, inefficient OPs have a wide variety of efficiency ratings, particularly with WRDD model scores ranging from −205.37 to 0.98 against those of the SBM model which goes from 0 to 0.96. On the one hand, the WRDD model has a weighted additive objective function (that weights the average percentual inefficiency of the inputs and outputs factors) whose inefficiency scores can be higher than one (i.e., it might be necessary for an increase higher than 100% on the outputs), and thus the inherent efficiency scores $\left(1-\beta_k^R\right)$ can be negative.

Table 1 Descriptive statistics of the inputs and outputs

Statistics	Total eligible spending	Number of operations	Total eligible costs decided
Mean	15,861,300	409	28,169,468
Standard deviation	38,520,025	1,068	63,497,428
Minimum	68,486	1	251,294
Maximum	237,904,467	5,457	311,154,920
Count	51	51	51

Source Authors’ own elaboration. Data available online at: https://cohesiondata.ec.europa.eu/2014-2020-Categorisation/ESIF-2014-2020-categorisation-ERDF-ESF-CF-planned-/3kkx-ekfq. (Accessed 30th March 2022)

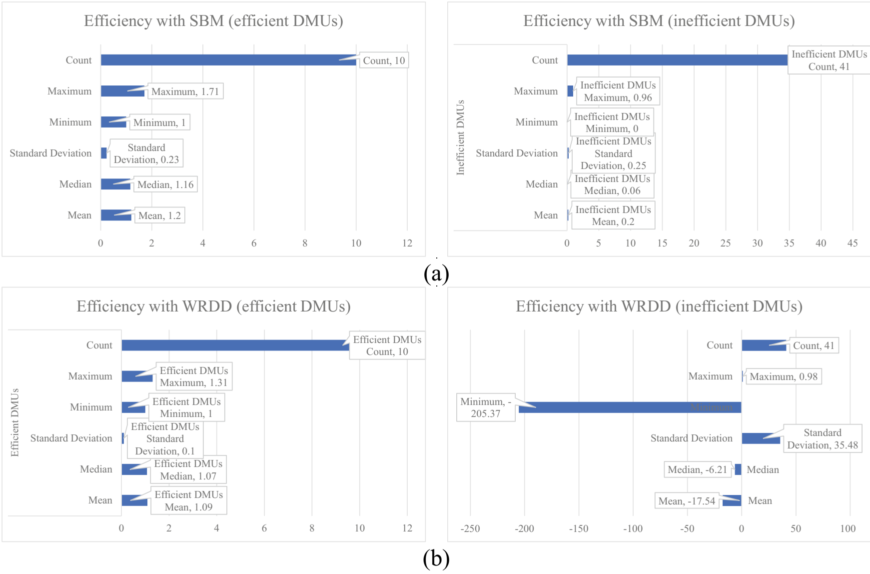


Fig. 1 Descriptive statistics of the results obtained with the SBM and WRDD models. *Source* Authors' own elaboration

Overall, these findings show the highest discriminatory power of inefficient OPs in the WRDD model.

The top four OPs elected as benchmarks are "Extremadura—ERDF" (27 and 29, with SBM and WRDD models, respectively), "País Vasco—ERDF" (26 and 24, with SBM and WRDD models, respectively), "Provence-Alpes-Côte d'Azur—ERDF/ESF/YEI" (20 with both models) and "Multi-regional Spain—ERDF" (14 and 21, with SBM and WRDD models, respectively)—see Fig. 2. Remarkably, each of these regions is located in MS that are among the biggest players in ICT support for SMEs (Pellegrin et al., 2018). Furthermore, the results of the Spanish regional OPs are in line with the conclusions of Ruiz-Rodríguez et al. (2018). For these researchers, Spanish regions have an intermediate or greater degree of digital innovation than their EU peers, as well as a reduced digital divide (DD) (i.e., a smaller disparity in firm digitization) than that of the other European MS. Interestingly, in the research reported by Ruiz-Rodríguez et al. (2018), Greece and Bulgaria were identified as the countries whose enterprises were last placed concerning digital adoption (based on 2015 data), highlighting the measures undertaken by these MS in the implementation of ICT by SMEs in the latest program time frame.

Both the SBM and the WRDD models also deliver information on the required changes that inputs and outputs should undergo to turn inefficient OPs into efficient ones (Fig. 3).

The 'number of operations supported' has the largest potential for improvement in both models (313% and 376%, according to the SBM and WRDD models, respectively), while 'eligible cost decided' necessitates a higher decrease with the SBM

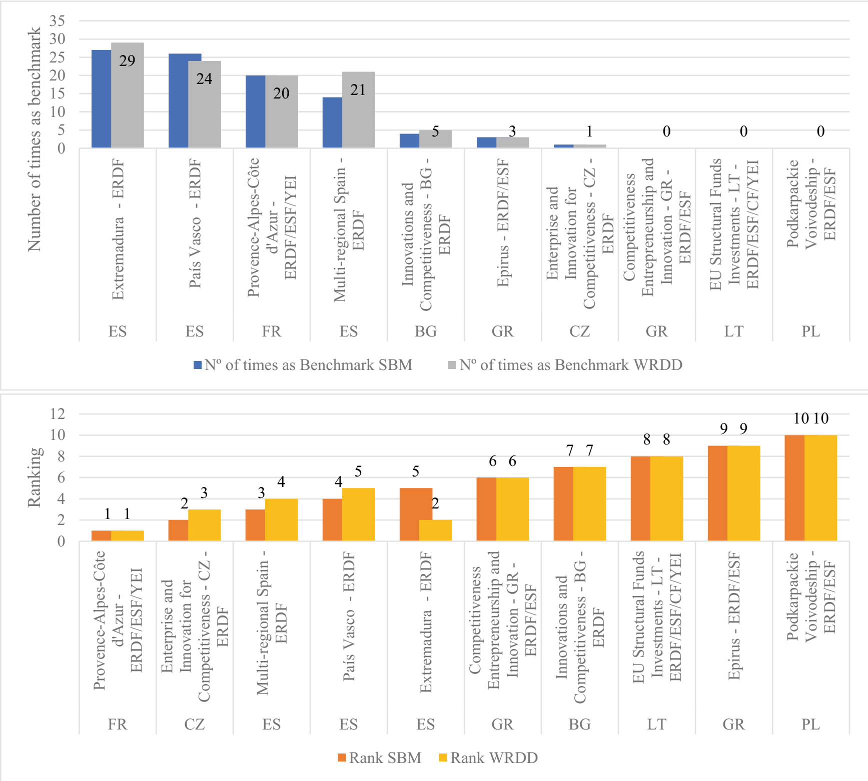


Fig. 2 Characteristics of efficient OPs according to the SBM and WRDD models. *Source* Authors' own elaboration

model (−22% against −5% with the WRDD model) and 'eligible spending' necessitates smaller changes with this same model (2% against 17% with the WRDD model)—see Fig. 3. Overall, the WRDD ends up being more demanding about the outputs whereas the SBM model is more demanding regarding the input. Finally, with the decomposition of inefficiency (with the WRDD model) it is also possible to conclude that the 'Number of operations' supported has an average effect on the inefficiency of 93%, while 'Total eligible spending' only has an impact of about 6%.

4.1 Robustness Study

Deterministic values are used both for inputs and outputs in conventional DEA approaches. However, the values considered to represent the input and output data are sometimes uncertain. Usually, the original DEA model is converted into two models, thus allowing to attain the upper and lower bounds of the efficiency scores. In the

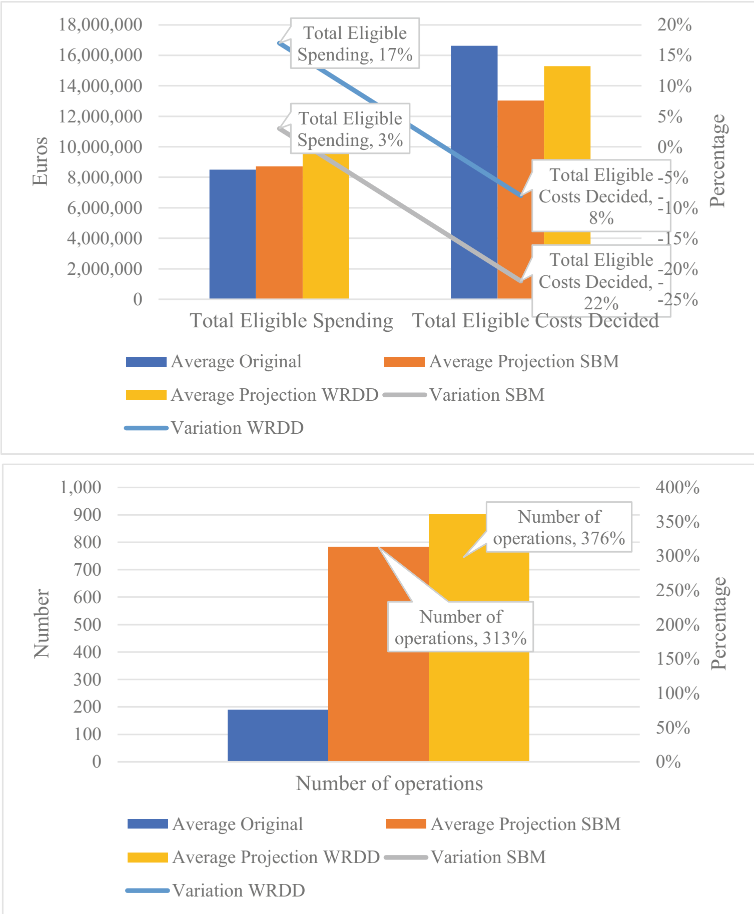


Fig. 3 Improvement potential for the OPs with the SBM and WRDD models. *Source* Authors' own elaboration

first model, all DMUs' outputs are raised while their inputs are lowered, except for the DMU under examination (i.e., DMU_k worsens its efficiency performance while the remaining DMUs improve their efficiency performance). In the second scenario, the opposite case occurs (Henriques & Marcenaro, 2021). The robustness assessment was conducted by employing data perturbations of 5% and 10% on the original data.

4.1.1 Robustness Study

According to our findings in Figs. 4a and b, only four OPs managed to hold efficiency for both data change scenarios according to both DEA models: "Enterprise and Innovation for Competitiveness—CZ—ERDF" (1st in terms of robustness with the SBM and 3rd with the WRDD models, respectively), "Provence-Alpes-Côte d'Azur—ERDF/ESF/YEI" (2nd in terms of robustness with both models), "Multi-regional Spain—ERDF" (3rd in terms of robustness with the SBM and 4th with the WRDD models, respectively), and "Extremadura—ERDF" (4th in terms of robustness with the SBM and 1st with the WRDD models, respectively). Additionally, while with the WRDD model 41% of the OPs are robustly inefficient, with the SBM model this value rises to 55%. Ultimately, these data indicate a weak utilization of ERDF in SMEs' ICT adoption. These findings support those of Pellegrin et al. (2018), who established that the EU lags behind its competitors (United States, Japan, and South Korea) in respect of ICT usage and connectivity, especially among SMEs.

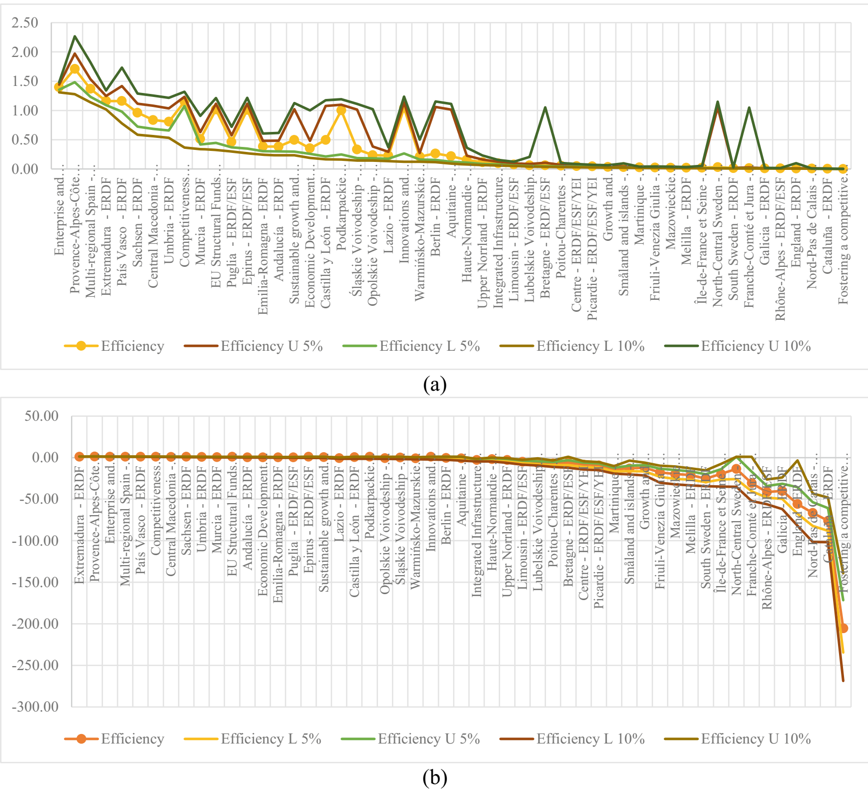


Fig. 4 Robustness analysis per OP with the SBM **a** and WRDD **b** models. *Source* Authors' own elaboration

5 Conclusions and Further Research

The main objective of this paper is to contrast the efficiency reached in the procedural efficiency of 51 OPs related to ICT assistance in SMEs from 16 EU MS according to two DEA models. We employ both the SBM and the WRDD models, which are both non-radial and non-oriented, to evaluate if the results obtained were robust.

Our main conclusions are provided next.

RQ1: "What indicators restrict the successful use of ERDF allocated to enhance ICT adoption in EU SMEs?"

The 'number of operations supported' has the greatest room to improve in both models (313% and 376%, as per SBM and WRDD models, respectively), whereas 'eligible spending decided' usually requires a significant decline with the SBM model (−22% versus −5% with the WRDD model) and 'eligible spending' requires minor changes with the very same model (2% versus 17% with the WRDD model). Essentially, the WRDD model is much more demanding of outputs, while the SBM model is much more demanding of inputs. Finally, based on the breakdown of inefficiency (using the WRDD model), it is reasonable to determine that the 'Number of operations' supported has a 93% average influence on inefficiency, whilst 'Total eligible spending' only has a 3% impact.

RQ2: "What OPs are generally elected as benchmarks across the assessed time frame?"

The four leading benchmarks are "Extremadura—ERDF" (27 and 29, using SBM and WRDD models, respectively), "País Vasco—ERDF" (26 and 24, using SBM and WRDD models, respectively), "Provence-Alpes-Côte d'Azur—ERDF/ESF/YEI" (20 using both models), and "Multi-regional Spain—ERDF" (14 and 21, using SBM and WRDD models, respectively). Surprisingly, every one of these regions is situated in MS that have a major participation in ICT assistance for SMEs (Pellegrin et al., 2018). In this regard, it is worth noting that two of these OPs are in Spain (Pellegrini et al., 2018), where the use of "vouchers" has proven to be an effective means of engaging SMEs and giving them the help that is simple to administer and specific to their requirements.

RQ3: "What OPs demonstrate more resilience in terms of their efficiency classification in the context of changes in the indicators used?"

Just four OPs continued to preserve efficiency for both data changes under both DEA models: "Enterprise and Innovation for Competitiveness—CZ—ERDF" (1st in terms of robustness with the SBM and 3rd with the WRDD models, respectively), "Provence-Alpes-Côte d'Azur—ERDF/ESF/YEI" (2nd in terms of robustness with both models), "Multi-regional Spain—(4th in terms of robustness with the SBM and 1st with the WRDD models, respectively). Furthermore, whereas the WRDD model has 41% of the OPs being robustly inefficient, the SBM model has 55%.

All in all, the results obtained with both models are generally consistent. In the end, it can be stated that SMEs' recourse to ESIF (particularly ERDF) is restricted. These findings might be related to their lack of organizational knowledge to interact with the numerous technicalities involved in applying for and executing ERDF operations. As regards ICT, this challenge becomes more urgent when compared to 'conventional' SME operations. As a result, operations in a sector known for sudden growth, such as ICT, need greater flexibility and competence. Therefore, MA should seek ways to give extra support that simplify operations while meeting the needs of SMEs.

Additionally, our research highlights the shortage of measures for measuring the performance of ESIF funding committed to ICT help in SMEs. Future work should address the impact of contextual variables on the outcomes herein obtained to understand the true reasons behind these poor results.

Acknowledgements This work has been funded by European Regional Development Fund in the framework of Portugal 2020—Programa Operacional Assistência Técnica (POAT 2020), under project POAT-01-6177-FEDER-000044 ADEPT: Avaliação de Políticas de Intervenção Co-financiadas em Empresas. INESC Coimbra and CeBER are supported by the Portuguese Foundation for Science and Technology funds through Projects UID/MULTI/00308/2020 and UIDB/05037/2020, respectively.

References

Bach, M. P., Zoroja, J., & Vukšić, V. B. (2013). Determinants of firms' digital divide: A review of recent research. *Procedia Technology, 9*, 120–128. https://doi.org/10.1016/j.protcy.2013.12.013

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*, 429–444. https://doi.org/10.1016/0377-2217(79)90229-7

Chen, P. C., Yu, M. M., Chang, C. C., Hsu, S. H., & Managi, S. (2015). Nonradial directional performance measurement with undesirable outputs: An application to OECD and Non-OECD countries. *International Journal of Information Technology & Decision Making, 14*(03), 481–520. https://doi.org/10.1142/S0219622015500091

Consoli, D. (2012). Literature analysis on determinant factors and the impact of ICT in SMEs. *Procedia-Social and Behavioral Sciences, 62*, 93–97. https://doi.org/10.1016/j.sbspro.2012.09.016

Henriques, C. O., & Marcenaro-Gutierrez, O. D. (2021). Efficiency of secondary schools in Portugal: A novel DEA hybrid approach. *Socio-Economic Planning Sciences*, *74*, 100954. S0038012120307916 100954 10.1016/j.seps.2020.100954.

Isensee, C., Teuteberg, F., Griese, K. M., & Topi, C. (2020). The relationship between organizational culture, sustainability, and digitalization in SMEs: A systematic review. *Journal of Cleaner Production, 275*, 122944. https://doi.org/10.1016/j.jclepro.2020.122944

Kleibrink, A., Niehaves, B., Palop, P., Sörvik, J., & Thapa, B. E. (2015). Regional ICT innovation in the European Union: Prioritization and performance (2008–2012). *Journal of the Knowledge Economy, 6*(2), 320–333. https://doi.org/10.1007/s13132-015-0240-0

Lehner, F., & Sundby, M.W. (2018). ICT skills and competencies for SMEs: Results from a structured literature analysis on the individual level. In: Harteis, C. (eds) The impact of digitalization in the workplace. Professional and practice-based learning (Vol. 21). Cham: Springer. https://doi.org/10.1007/978-3-319-63257-5_5

Mbuyisa, B., & Leonard, A. (2017). The role of ICT use in SMEs towards poverty reduction: A systematic literature review. *Journal of International Development, 29*(2), 159–197. https://doi.org/10.1002/jid.3258

Oberländer, M., Beinicke, A., & Bipp, T. (2020). Digital competencies: A review of the literature and applications in the workplace. *Computers & Education, 146*, 103752. https://doi.org/10.1016/j.compedu.2019.103752

Pellegrin, J., Colnot, L., Arendt, Ł., Bisaschi, L., Catalano, G., Martinaitis, Ž., & Micheletti, G. (2018). Research for REGI committee-digital agenda and cohesion policy. European Parliament. Retrieved 30/05/2022, https://www.europarl.europa.eu/RegData/etudes/STUD/2018/617485/IPOL_STU(2018)617485_EN.pdf.

Ramdani, B., Raja, S., & Kayumova, M. (2022). Digital innovation in SMEs: A systematic review, synthesis and research agenda. *Information Technology for Development, 28*(1), 56–80. https://doi.org/10.1080/02681102.2021.1893148

Reggi, L., & Gil-Garcia, J. R. (2021). Addressing territorial digital divides through ICT strategies: Are investment decisions consistent with local needs? *Government Information Quarterly, 38*(2), 101562. https://doi.org/10.1016/j.giq.2020.101562

Reggi, L., & Scicchitano, S. (2014). Are EU regional digital strategies evidence-based? An analysis of the allocation of 2007–13 Structural Funds. *Telecommunications Policy, 38*(5–6), 530–538. https://doi.org/10.1016/j.telpol.2013.12.007

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432–448). Routledge.

Ruiz-Rodríguez, F., Lucendo-Monedero, A. L., & González-Relaño, R. (2018). Measurement and characterisation of the Digital Divide of Spanish regions at enterprise level. A comparative analysis with the European context. Telecommunications Policy, *42*(3), 187–211.

Tarutė, A., & Gatautis, R. (2014). ICT impact on SMEs performance. *Procedia-Social and Behavioral Sciences, 110*, 1218–1225. https://doi.org/10.1016/j.sbspro.2013.12.968

Taylor, P. (2015). The importance of information and communication technologies (ICTs): An integration of the extant literature on ICT adoption in small and medium enterprises. International Journal of Economics, Commerce and Management, 3(5).

Thrassou, A., Uzunboylu, N., Vrontis, D., & Christofi, M. (2020). Digitalization of SMEs: A review of opportunities and challenges. In: Thrassou, A., Vrontis, D., Weber, Y., Shams, S. M. R., Tsoukatos, E. (Eds.) The changing role of SMEs in global business. Palgrave Studies in Cross-disciplinary Business Research, In Association with EuroMed Academy of Business. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-45835-5_9.

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research, 130*(3), 498–509. https://doi.org/10.1016/S0377-2217(99)00407-5

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). Processes of technological innovation. Lexington books.

Are ERDF Devoted to Boosting ICT in SMEs Inefficient? Further Insights Through the Joint Use of DEA with SFA Models

Carla Henriques and Clara Viseu

Abstract We employ a three-stage data envelopment analysis (DEA) model coupled with Stochastic Frontier Analysis (SFA) by using data made freely available by the European Commission, to evaluate the procedural efficiency of 51 OPs from 16 countries committed to fostering the adoption of Information and Communication Technologies (ICT) in small mid-sized enterprises (SMEs). We depart from the results obtained in the previous Chapter with two DEA models, specifically the Slack Based-Measure (SBM) and the Weighted Russel Directional Distance (WRDD) model. Firstly, we adjust the input and output factors through the SFA by removing the influence of environmental factors and statistical noise. Secondly, we instantiate the previous DEA models with adjusted factors, to compute new efficiency factors. All in all, we observed that by removing these contextual effects, nearly 27% of the OPs (14) vs. 30% of the OPs (16) using the SBM and the WRDD approaches, respectively, achieved efficient procedural outcomes, compared to 20 percent (10) without the consideration of these factors. The OP "Multi-regional Spain - ERDF" is commonly considered a benchmark regardless of the model and contextual environment. The 'number of operations supported' is the measure that necessitates more consideration, with or without the elimination of environmental factors, irrespective of the DEA model used. Our findings imply that more developed areas with a higher number of ICT specialists tend to have lower use of ERDF funds committed to promoting ICT adoption in SMEs. These findings might be attributable to administrative practices and SMEs' failure to manage the complicated technicalities involved in submitting

C. Henriques (✉) · C. Viseu
Polytechnic of Coimbra, Coimbra Business School, 3045-601 Coimbra, Portugal
e-mail: chenriques@iscac.pt

C. Henriques
INESC Coimbra—Department of Electrical and Computer Engineering (DEEC), University of Coimbra, Polo 2, 3030-290 Coimbra, Portugal

Centre for Business and Economics Research (CeBER), Faculty of Economics, University of Coimbra, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

C. Viseu
Centro de Matemática e Aplicações (CMA-UBI), Universidade da Beira Interior, 6200-001 Covilhã, Portugal

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_3

and executing European Regional Development Fund (ERDF) project proposals. As a result, it is critical to provide the extra help that reduces managerial requirements while also meeting the demands of SMEs.

Keywords ICT · SMEs · SBM model · SFA · ERDF

1 Introduction

Given the advantages and possibilities that ICT-based technology may provide, and its rapid acceptance over the years, SMEs have been unable to strengthen its use (Akter et al., 2020; Haaker et al., 2021). Emerging innovations, particularly ICT, remain to pose challenges for businesses (Martin & Leurent, 2017; Oberländer et al., 2020; Vial, 2019). This might be attributable, in effect, to SMEs' scarce funds, equipment, and experience. In practice, many hurdles to SMEs' usage of ICT exist (Consoli, 2012): Economical, since significant investments are required and money is difficult to obtain; infrastructural, specifically owing to power pricing, bandwidth, and secure Internet connection; organizational, mostly shortage of skilled people; and technological, because technology advance involves careful preparation. Another issue impeding SMEs' usage of ICT is a total lack of awareness of the capabilities and repercussions of digitalization (Horváth & Szabó, 2019; Lehner & Sundby, 2018). SMEs are concerned that if they do not pursue digitization, they would lose profitability (Li et al., 2018; Ulas, 2019), but managers are hesitant to embrace it because they are unaware of how to integrate them into the business (Lehner & Sundby, 2018; Reis et al., 2018).

Governments should establish measures to lower the digital divide (DD), offer affordable network connectivity, and invest in education to encourage ICT usage in SMEs (Horváth & Szabó, 2019; Lehner & Sundby, 2018; Reis et al., 2018).

Ex-post assessments are frequently utilized in research evaluating EU structural funding awarded to ICT (Kleibrink et al., 2015; Reggi & Scicchitano, 2014). There is additional evidence that supports an ex-ante evaluation of the factors influencing funding distribution amongst different ICT activities (Reggi & Gil-Garcia, 2021). The research conducted in the previous Chapter contrasted the implementation of OPs connected to ICT policies with their equivalents throughout the programming period, illustrating the modifications that should be performed to render an inefficient OP efficient. Nonetheless, no viable reason for the influence of contextual variables on the outcomes obtained has been investigated. Consequently, the objective of this research is to contribute to the literature by employing a systematic framework that helps Management Authorities (MA) to assess the reasons behind the inefficiency of the implementation of the OPs committed to supporting ICT adoption by SMEs employing two non-parametric methods, namely the SBM and WRDD models, in conjunction with SFA.

Below are the key questions that this work intends to answer:

RQ1: "Which environmental factors have the largest effect on the OPs' inefficiency?".

RQ2: "How does efficiency differ when contextual factors and statistical noise are removed?".

Hereunder is the structure for this article. Section 2 describes the fundamental premises behind the analytical techniques proposed to evaluate the implementation of the OPs under consideration. Section 3 discusses the main reasons for selecting the indicators, and the contextual factors used for the evaluation, but also the descriptive statistics on these data. The main results are discussed in Sect. 4. Section 5 summarizes the interesting conclusions, examines possible political implications, identifies critical weaknesses, and suggests future research directions.

2 Methodology

One shortcoming of the DEA approach is that it does not account for the effect of contextual factors and random errors when evaluating efficiency. Fried et al. (2002) proposed a three-stage DEA model consequently. To begin, the DEA model is used to calculate the efficiency scores of each DMU, as well as the necessary changes to the input and output components to convert inefficient DMUs into efficient DMUs (i.e., the slacks). Second, the slacks are cate;gorized into three parts: contextual issues, poor management, and statistical noise. The slacks are the dependent variables, while the contextual factors are the independent variables. The aim is to minimize the influence of contextual factors and random errors. SFA is then used to adjust the input and output factors (Aigner et al., 1977; Meeusen & van den Broeck, 1977).

Let n be the set of DMUs $(\mathrm{DMU}_1, \mathrm{DMU}_2, \ldots, \mathrm{DMU}_\mathrm{n})$ with $X = \left[x_{ij},\ i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n\right]$ the $(m \times n)$ matrix of *inputs*, $Y = \left[y_{rj},\ r = 1, 2, \ldots, s,\ j = 1, 2, \ldots, n\right]$ the vector of *outputs* $(s \times n)$ and the rows of these matrices for DMU_k are, respectively, $\boldsymbol{x}_k^T$ and $\boldsymbol{y}_k^T$, where $^{\mathrm{T}}$ is the transpose of a vector. Also, assume a Variable Returns to Scale (VRS) technology with the imposition of $\sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \geq 0(\forall_j)$.

The SBM model

$$\underset{\lambda, s^-, s^+}{Min} \rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} s_i^-/x_{ik}}{1 + \frac{1}{s}\sum_{r=1}^{s} s_r^+/y_{rk}}$$

s.t.

$$\begin{aligned} x_{ik} &= \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-,\ i = 1, \ldots, m \\ y_{rk} &= \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, r = 1, \ldots, s \\ &\lambda_j \geq 0,\ \ j = 1,\ ...,\ n \\ &s_i^- \geq 0,\ i = 1, \ldots, m \\ &s_i^+ \geq 0, r = 1, \ldots, s \end{aligned} \tag{1}$$

The WRDDM

$$\max\left(w_y\left(\sum_r \varpi_y^r \frac{s_r^{+'}}{g_{yr}}\right) + w_x\left(\sum_i \varpi_x^i \frac{s_i^{-'}}{g_{xi}}\right)\right)$$

$$\begin{aligned} &\text{s.t.} \sum_{j=1}^{n} \lambda_j y_{rj} = y_{ro} + s_r^{+'},\ r = 1, \ldots, s, \\ &\sum_{j=1}^{n} \lambda_j x_{ij} = x_{io} - s_i^{-'}, i = 1, \ldots, m, \\ &\sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \ldots, n, \\ &s_r^{+'} \geq 0(\forall_r), s_i^{-'} \geq 0(\forall_i) \end{aligned} \tag{2}$$

The directional vectors $\boldsymbol{g}_x$ and $\boldsymbol{g}_y$ are such that $(-\boldsymbol{g}_x, \boldsymbol{g}_y) = (-\boldsymbol{x}_k, \boldsymbol{y}_k)$

The SBM model	The WRDDM
where the value of $0 < \rho < 1$, and DMU_k is SBM-efficient if $\rho^* = 1$, meaning that the slacks (s_i^- and s_i^+) are null for all the inputs and outputs	w_y and w_x, with $w_y + w_x = 1$, are weights that assign the importance of the outputs and inputs. The importance of the inefficiencies of every factor is defined such that $\sum_{r\epsilon O} \varpi_y^r = 1, \sum_{i\epsilon I} \varpi_x^i = 1$

Every output[1] slack obtained for j inefficient DMU ($j = 1, \ldots, p$) is:

$$s_{rj} = f\left(Z_j, \beta^r\right) + v_{rj} + u_{rj}, r = 1, \ldots, s;\ j = 1, \ldots, p, \tag{3}$$

where s_{rj} is the slack value of output r of DMUj, $f\left(Z_j, \beta^r\right)$ is the deterministic feasible slack frontier, and β^r denotes the coefficients associated with the contextual factors. The term $v_{rj} + u_{rj}$ is the mixed error, v_{rj} is the statistical noise and u_{rj} is management inefficiency. Usually, it is assumed that $v_{rj}\ N\left(0; \sigma_v^2\right)$ and $u_{rj}\ N^+\left(\mu^i; \sigma_u^2\right)$, with v_{rj} and u_{rj} independent variables.

Let $\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$. If γ is close 1, it indicates that management factors are the main responsible for the adjustment needed to achieve efficiency. If $\gamma$is near 0, most of the adjustment needed to achieve efficiency is linked to statistical noise.

The modified output slacks are then calculated by decomposing the mixed error. In line with Jondrow et al. (1982), the management inefficiency is calculated as follows:

$$E\left(u_{ij}|u_{ij} + v_{ij}\right) = \frac{\sigma\delta}{1 + \delta^2}\left[\frac{\varphi\left(\frac{\varepsilon_j\delta}{\sigma}\right)}{\emptyset\left(\frac{\varepsilon_j\delta}{\sigma}\right)} + \frac{\varepsilon_j\delta}{\sigma}\right] \tag{4}$$

where $\delta = \frac{\sigma_u}{\sigma_v}$, $\varepsilon_j = v_{ij} + u_{ij,}$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$, φ and;$\emptyset$ are, respectively, the density and distribution functions of the standard normal distribution. Thus, the random error term can be obtained as:

$$E\left(v_{ij}|u_{ij} + v_{ij}\right) = s_{ij} - f\left(Z_j, \beta^i\right) - E\left(u_{ij}|u_{ij} + v_{ij}\right). \tag{5}$$

Based on the three-stage method of Fried et al. (2002), at the first stage, the slacks are computed through the SBM and WRDD models. At the second stage, the output variables of each DMU are modified according to the SFA results by removing the significant contextual effects and statistical noises.

The adjusted outputs are obtained as (Avkiran & Rowlands, 2008):

$$y_{rj}^A = y_{rj} + \left[f\left(Z_j, \beta^r\right) - \min_r\left\{f\left(Z_j, \beta^r\right)\right\}\right] + \left[v_{rj} - \min_r\left\{v_{rj}\right\}\right],\ r = 1, \ldots, s. \tag{6}$$

[1] We only describe output adjustments since these were the only factors that requeired further attention.

Finally, at the third stage, the efficiency scores are obtained with the modified output factors.

3 Data

The input and output factors used for assessing the efficiency of the implementation of ERDF engagement in ICT adoption in SMEs were obtained from the list of indicators formally reported to the EU.[2]

Because of the missing data on ICT adoption at the enterprise level from traditional databases per NUTS2 region (Billon et al., 2016, 2017; Reggi & Gil-Garcia, 2021), we employ metrics available from the Regional Innovation Scoreboard in 2021 as contextual/environmental variables (Hollanders, 2021). Additional statistical data were gathered from OECD statistics.[3]

3.1 Input and Output Factors

3.1.1 "Total Eligible Costs Decided" and "Total Eligible Spending"

The indicators employed to measure the efficiency of the OPs' absorption are "total eligible spending" and "total eligible costs decided.“ The first concerns qualified costs that have been documented and verified by a decision authority. Consequently, this component is employed as an output since the more the value assigned to it, the greater the financial implementation of each operation. The second is regarded as an input since it relates to the financial resources that are assigned to the programs selected for finance, which must be maintained as low as possible.

3.1.2 Number of Operations Supported

The "number of operations supported" alludes to the number of projects financed by the ERDF. The greater the number of projects funded, the greater the prospect of raising ICT usage. Therefore, this indicator is an output.

Further information on these data is available in the previous Chapter.

[2] Available online: https://cohesiondata.ec.europa.eu/2014-2020-Categorisation/ESIF-2014-2020-categorisation-ERDF-ESF-CF-planned-/3kkx-ekfq (accessed 30th March 2022).

[3] Available online: https://stats.oecd.org/Index.aspx?DataSetCode=REGION_ECONOM (accessed 30th March 2022).

3.2 Environmental Factors

As a measure of economic growth, we utilized regional GDP at purchasing power parity per capita (GDPPPpc) as an exogenous variable (Billon et al., 2016, 2017; Reggi & Gil-Garcia, 2021). Besides, ICT is more successfully implemented in affluent locations, as suggested by Neokosmidis et al. (2015).

We additionally used as an environmental variable the proportion of the population aged 25–34 who have completed a college education, as data suggests a positive association between educational achievement and ICT usage (Billon et al., 2016, 2017). Several factors have been addressed in the context of ICT adoption to justify this seeming advantageous relationship. On the one hand, education teaches the qualifications needed to use and profit from the use of ICT. Workers, on the other hand, are expected to become more acquainted with ICT (Billon et al., 2017). Since it has been acknowledged that dissemination of ICT in EU regions is positively linked with Research and Development (R&D) expenditures (Billon et al., 2016, 2017; Giotopoulos et al., 2017), we used R&D investment in firms as a percentage of GDP and the number of SMEs attempting to bring new to market products as a percentage of all SMEs as potential explanatory factors of ICT adoption.

Additionally, since a firm's abilities are regarded as important innovation elements influencing user and ICT acceptance (Giotopoulos et al., 2017), we considered workers with basic digital skills as a proportion of total SMEs' workers. Finally, the percentage of ICT specialists as a percentage of overall SMEs' labor was also used, i.e., workers whose primary occupation is ICT and who can manage a broad variety of duties linked to computers (Ruiz-Rodríguez et al., 2018).

Figure 1 shows the normalized scores, which vary from 0 to 1, except for the GDPPPpc, which was quantified by an index value.

4 Discussion of Results

The SFA regression models were run through the R software, version 4.0.5 (RStudio Team, 2021), namely, the sfaR package version 0.1.1 (Dakpo et al., 2022). The results thus computed with both models are depicted in Table 1.

The results of γ in both models are close to one and statistically significant (1%), indicating that management failures were the major reason behind the attained inefficiency scores. We employed SFA to exclude the effects of contextual factors and random errors to obtain neutral efficiency estimates. The regression coefficients are all significant (1%), showing that the given environmental variables have a substantial effect on the slacks calculated.

Growth in both the proportion of ICT specialists and GDPPPpc, according to Table 1, adds to a higher required rise in "total eligible spending", whereas the remaining factors have a negative influence on this slack. These findings show that wealthier areas and a greater number of ICT professionals might not always indicate

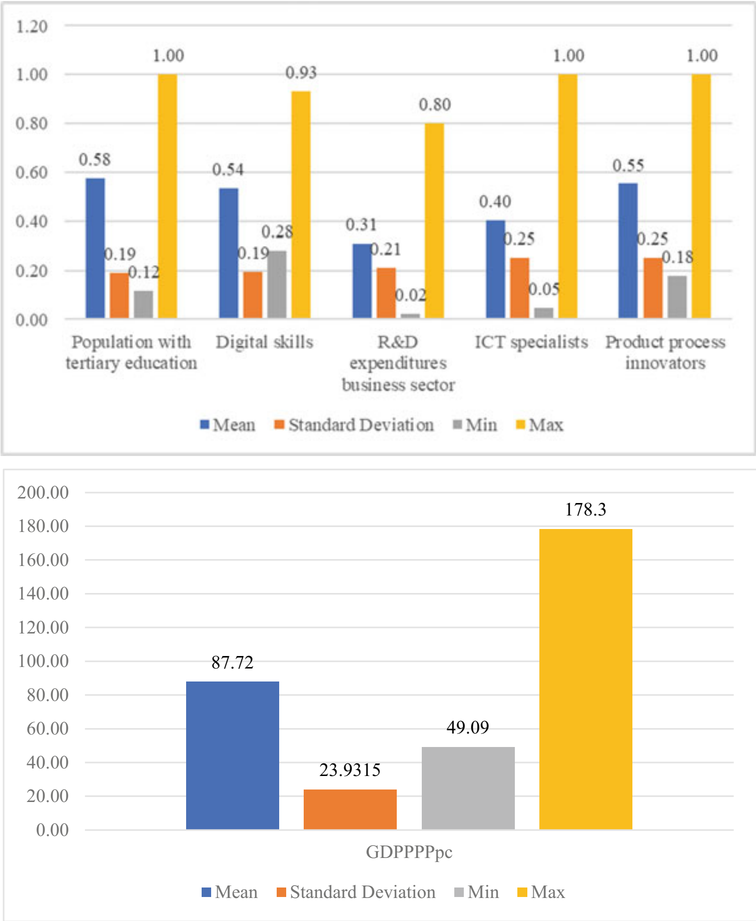


Fig. 1 Descriptive Statistics of the contextual variables. *Source* Authors' own elaboration

a higher rate of implementation of ERDF aimed at boosting ICT in SMEs. Following these findings, Bukvić et al. (2021) identified ERDF underuse by ICT Croatian SMEs from 2014 to 2020. Their study claimed that the difficulties and time required to implement, design, and assess the projects might explain these results (Bukvić et al., 2021). Moreover, Martinez-Cillero et al. (2020) revealed that SMEs' expenditures are lower than traditional economic models would anticipate, implying that these businesses are highly susceptible to funding issues. Another reason for these findings might be that these SMEs make use of alternative funding sources (Pellegrin et al., 2018).

Concerning the requirement to improve the 'number of operations supported', we also found that this aspect tends to rise as digital skills and ICT specialists grow, but

Table 1 Results with SFA for both models

Variables	Slacks SBM		Slacks WRDD	
	Total Eligible Spending	Number of Operations	Total Eligible Spending	Number of Operations
Constant	−242,050***	237.20***	−696,940***	248.61***
Population with tertiary education	−890,650***	–	−4,007,800***	–
Digital skills	−890,970***	195.99***	−3,279,600***	205.16***
ICT specialists	1,417,700***	135.17***	6,634,300***	143.90***
Product process innovators	–	−73.17***	−3,805,600***	−83.38***
$GDPPPP_{pc}$	1286***	−3.39***	15,499***	−3.52***
Sigma-squared	8.91E+11***	8.11E+05***	2.85E+13***	1.08E+06***
Gamma	0.98**	0.99**	0.97**	0.99**
Log-likelihood function	−593.83	−308.67	−672.65	−331.54

** The model coefficients are statistically significant, at the 5% level of significance (p-value < 0.05)
*** The model coefficients are statistically significant, at the 1% level of significance (p-value < 0.01)
Source Authors' own elaboration

it begins to decrease when the percentage of SMEs with products process innovations and GDPPPPpc grow. These data demonstrate once again that a high proportion of ICT skills/specialists does not imply an appropriate 'number of operations supported.' Areas with a greater $GDPPPP_{PC}$ and a greater number of SMEs that are more receptive to process innovation, on the other hand, may not always need to apply for further ERDF initiatives since they are more efficient in obtaining funding.

Figure 2 shows that efficient OPs reduced their variance in terms of performance (the standard deviation is now 0.15 and 0.05 against 0.23 and 0.10 with the SBM and WRDD models, respectively). Furthermore, the efficiency ratings are restricted within [1.00, 1.49] and [1.00, 1.18] using the SBM and WRDD models, respectively. Moreover, inefficient OPs reduce the variability of their technical efficiency (with a standard deviation of 0.21 and 0.05 as opposed to the prior 0.25 and 35.48 with the SBM and WRDD models, respectively) and significantly boost their efficiency (underlining the importance of the contextual factors).

Figure 2 shows that about 27 and 31 percent of the OPs attained procedural efficiency with the SBM and WRDD models, respectively, relative to the earlier 20 percent, i.e., 10 out of 51.

Figure 3 depicts the variation in OPs' technical efficiency with and without modified parameters for both models for the OPs that become efficient with the adjusted factors.

When contrasted to the first stage of the assessment, "Berlin - ERDF", "Haute-Normandie—ERDF/ESF/YEI", "Central Macedonia—ERDF/ESF", "Puglia—ERDF/ESF", "Melilla—ERDF", "Umbria—ERDF", "Sachsen—ERDF", "Upper

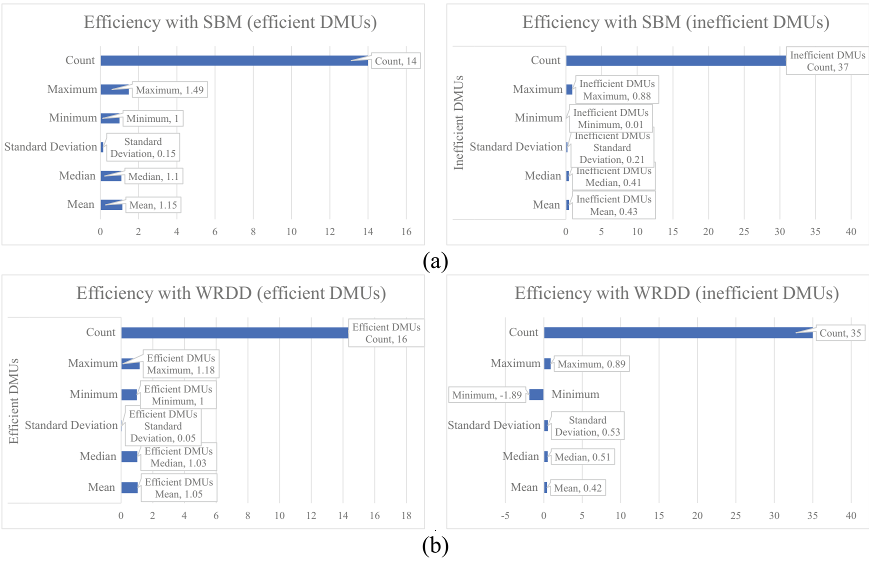


Fig. 2 Descriptive Statistics of the results obtained with adjusted factors for both SBM (**a**) and WRDD (**b**) models. *Source* Authors' own elaboration

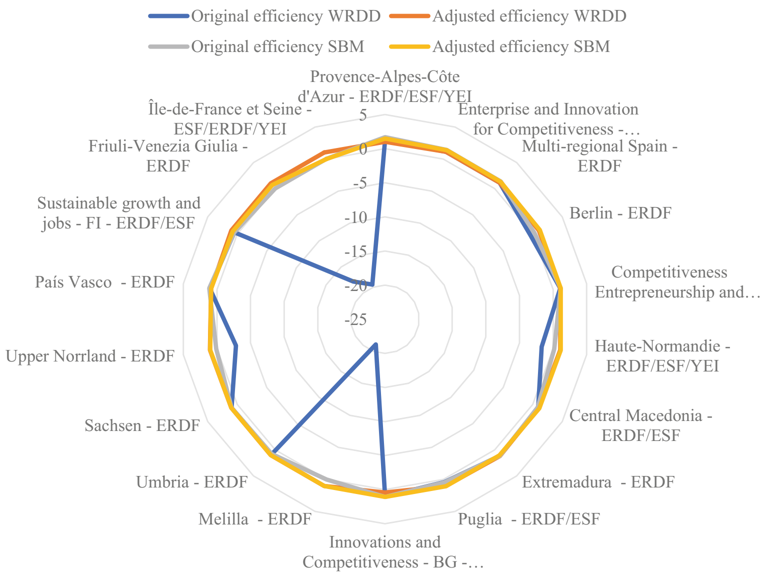


Fig. 3 OPs that that became efficient with adjusted factors with the SBM and WRDD models. *Source* Authors' own elaboration

Norrland—ERDF" all become efficient according with both models, whereas "País Vasco—ERDF" becomes inefficient according. These results indicate that the previous inefficiencies of these OPs were mainly impacted by their environmental factors.

From Table 2 it is visible that "Multi-regional Spain—ERDF" is the only efficient OP more immune to the model and adjustments considered being ranked in the 3rd place according to both models with the consideration of the adjusted factors. Now, "Central Macedonia—ERDF/ESF" (25—SBM and 12—WRDD), "Berlin—ERDF" (18—SBM and 10—WRDD), "Puglia—ERDF/ESF" (10—SBM and 2—WRDD) and " Multi-regional Spain—ERDF " (9—SBM and 18—WRDD) are the top 4 OPs more widely viewed as benchmarks—see Table 2.

Surprisingly, "Multi-regional Spain—ERDF" is one of the three leading efficient OPs, serving as a reference for best practices regardless of the model and the removal of environmental factors. Furthermore, two of the OPs that retain their efficiency despite the removal of environmental factors and the model used are from Spain. It is worth noting that MS in the Southern and Central and Eastern EU were the primary receivers of ICT and digital economy aid (Pellegrin et al., 2018). This is especially true for countries with efficient OPs, such as Spain, Greece, Bulgaria, and the Czech Republic (Pellegrin et al., 2018).

Finally, the enhancement that the 'number of operations supported' should undergo becomes substantially reduced falling from 313 and 376% to 111% and 141% using the SBM and WRDD, respectively—see Fig. 4.

5 Conclusions and Further Research

The objective of this paper is to evaluate the procedural efficiency of 51 OPs related to ICT assistance in SMEs from 16 EU MS, by including in the analysis the environmental factors that might impact efficiency outcomes. We presented a three-DEA modeling approach to achieve this goal. To begin, both the SBM and WRDD models are utilized to calculate the efficiency ratings of each OP. At this stage, pertinent information about the adjustments that should be made to alleviate any disparities between inefficient OPs and their benchmarks are collected.

The second phase includes the use of SFA to the slacks of inefficient OPs to update the inputs and outputs after removing contextual factors and statistical noises. At this stage, it is also feasible to comprehend how much contextual elements may influence the efficiency of ERDF implementation in distinct OPs devoted to increasing ICT adoption in SMEs, as well as the significance of management failures. Finally, the previously corrected components are employed in the SBM and WRDD models to obtain updated efficiency ratings.

Our important findings are listed below.

RQ1: "Which environmental factors have the largest effect on the OPs' inefficiency?".

Table 2 Characteristics of efficient OPs with adjusted factors according to both models

OP	Nº of times as Benchmark with adjustment—SBM	Nº of times as Benchmark without adjustment—SBM	Rank with adjustments—SBM	Rank without adjustment—SBM	Nº of times as Benchmark with adjustment—WRDD	Nº of times as Benchmark without adjustment—WRDD	Rank with adjustments—WRDD	Rank without adjustments—WRDD
Central Macedonia—ERDF/ESF	25	0	7	12	12	0	15	12
Berlin—ERDF	18	0	4	22	10	0	6	22
Puglia—ERDF/ESF	10	0	9	17	2	0	9	18
Multi-regional Spain—ERDF	9	14	3	3	18	21	3	4
Innovations and Competitiveness—BG—ERDF	7	4	10	7	0	5	43	7
Sachsen—ERDF	6	0	13	11	6	0	14	11
Haute-Normandie—ERDF/ESF/YEI	4	0	6	27	1	0	4	27
Umbria—ERDF	4	0	12	13	25	0	10	13
Extremadura—ERDF	2	27	8	5	5	29	1	2
Enterprise and Innovation for Competitiveness—CZ—ERDF	1	1	2	2	2	1	2	3
Provence-Alpes-Côte d'Azur—ERDF/ESF/YEI	0	20	1	1	0	20	16	1
Competitiveness Entrepreneurship and Innovation—GR—ERDF/ESF	0	0	5	6	0	0	5	6
Melilla—ERDF	0	0	11	43	0	0	12	43
Upper Norrland—ERDF	0	0	14	28	0	0	7	29
País Vasco—ERDF	0	0	15	4	0	24	19	5
Friuli-Venezia Giulia—ERDF	0	0	17	15	0	0	11	40
Sustainable growth and jobs—FI—ERDF/ESF	0	0	17	15	3	0	13	15
Île-de-France et Seine—ESF/ERDF/YEI	0	0	51	41	0	0	8	41

Source Authors' own elaboration

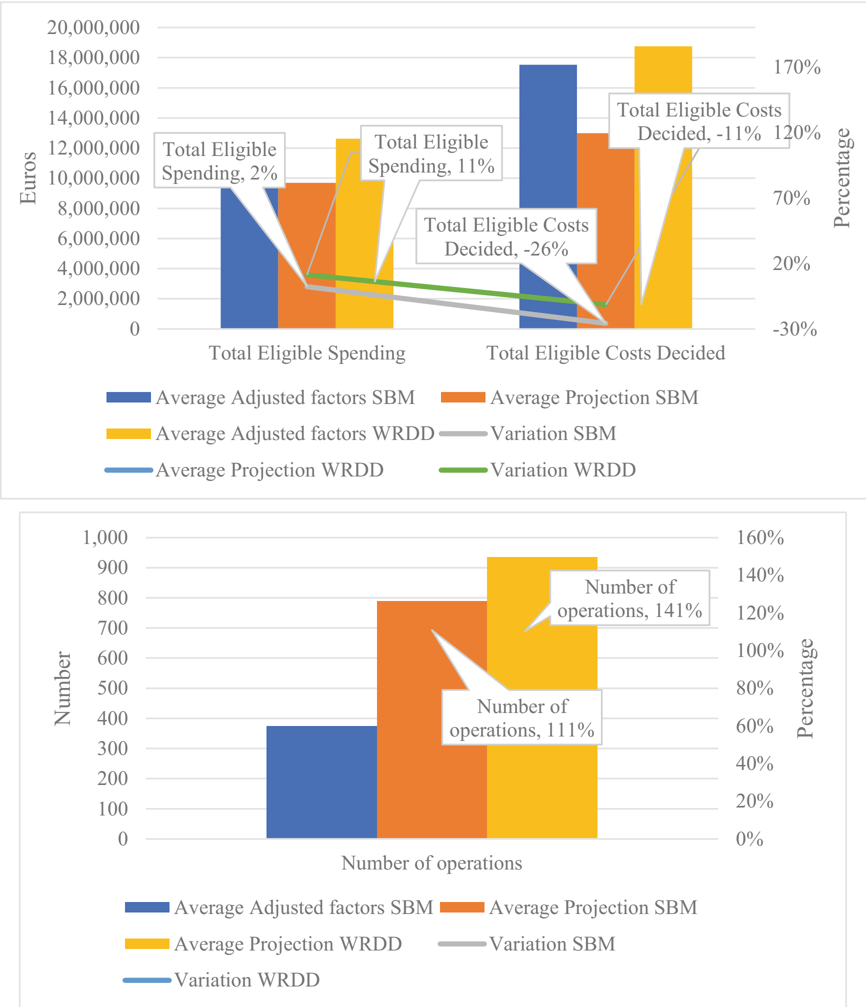


Fig. 4 Improvement potential for the OPs with the SBM and WRDD models. *Source* authors' own elaboration

Our findings show that richer regions with a larger number of ICT specialists underutilize ERDF funding designated for strengthening ICT in SMEs. Furthermore, a greater proportion of ICT skills/specialists equates to a lesser "number of operations supported." Richer locations and a bigger number of SMEs seeking product innovations, on the other hand, tend to be more efficient in receiving economic support. These results may be related to bureaucracy challenges in SMEs' conformity with EU protocols, financial mechanisms, and administrative procedures.

RQ2: "How does efficiency differ when contextual factors and statistical noise are removed?".

With the removal of the environmental factors, more than 27% and 30% of the OPs (14 and 16) achieved efficiency levels, according to the SBM and WRDD models, respectively, compared with the previous 20 percent (10), demonstrating the importance of environmental factors in efficiency assessment.

Consequently, it can be stated that SMEs' recourse to ESIF (particularly ERDF) is restricted since they lack the organizational ability to deal with the numerous procedures involved in applying for and executing ERDF projects. When it comes to ICT, this challenge becomes more urgent when compared to 'conventional' SME operations. As a result, operations in a sector known for rapid change, such as ICT, need greater flexibility and competence. As a result, MA must seek ways to give special support that simplify operations while meeting the needs of SMEs.

Additionally, our research highlights the shortage of measures for measuring the performance of ESIF funding committed to ICT help in SMEs. Finally, while this study revealed new insights and innovative ways for examining the efficiency of funding execution allocated to boosting ICT usage in EU SMEs, future research should focus on the economic repercussions of these OPs, which remains a difficult task.

Acknowledgements This work has been funded by European Regional Development Fund in the framework of Portugal 2020 – Programa Operacional Assistência Técnica (POAT 2020), under project POAT-01-6177-FEDER-000044 ADEPT: Avaliação de Políticas de Intervenção Co-financiadas em Empresas. INESC Coimbra and CeBER are supported by the Portuguese Foundation for Science and Technology funds through Projects UID/MULTI/00308/2020 and UIDB/05037/2020, respectively.

References

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics, 6*(1), 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

Akter, S., Michael, K., Uddin, M. R., McCarthy, G., & Rahman, M. (2020). Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*, 1–33. https://doi.org/10.1007/s10479-020-03620-w.

Avkiran, N. K., & Rowlands, T. (2008). How to better identify the true managerial performance: State of the art using DEA. *Omega, 36*(2), 317–324. https://doi.org/10.1016/j.omega.2006.01.002

Billon, M., Lera-Lopez, F., & Marco, R. (2016). ICT use by households and firms in the EU: Links and determinants from a multivariate perspective. *Review of World Economics, 152*(4), 629–654. https://doi.org/10.1007/s10290-016-0259-8

Billon, M., Marco, R., & Lera-Lopez, F. (2017). Innovation and ICT use by firms and households in the EU: A multivariate analysis of regional disparities. *Information Technology and People, 30*(2), 424–448. https://doi.org/10.1108/ITP-05-2015-0098

Bukvić, I. B., Babić, I. Đ, & Starčević, D. P. (2021). Study on the Utilization of National and EU Funds in Financing Capital Investments of ICT Companies. In 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 1282–1287). IEEE. https://doi.org/10.23919/MIPRO52101.2021.9597077

Consoli, D. (2012). Literature analysis on determinant factors and the impact of ICT in SMEs. *Procedia-Social and Behavioral Sciences, 62*, 93–97. https://doi.org/10.1016/j.sbspro.2012.09.016

Dakpo KH., Desjeux Y., & Latruffe L. (2022). sfaR: Stochastic Frontier Analysis using R. R package version 0.1.1.

Fried, H. O., Lovell, C. K., Schmidt, S. S., & Yaisawarng, S. (2002). Accounting for environmental effects and statistical noise in data envelopment analysis. *Journal of Productivity Analysis, 17*(1), 157–174. https://doi.org/10.1023/A:1013548723393

Giotopoulos, I., Kontolaimou, A., Korra, E., & Tsakanikas, A. (2017). What drives ICT adoption by SMEs? Evidence from a large-scale survey in Greece. *Journal of Business Research, 81*, 60–69. https://doi.org/10.1016/j.jbusres.2017.08.007

Haaker, T., Ly, P. T. M., Nguyen-Thanh, N., & Nguyen, H. T. H. (2021). Business model innovation through the application of the Internet-of-Things: A comparative analysis. *Journal of Business Research, 126*, 126–136. https://doi.org/10.1016/j.jbusres.2020.12.034

Hollanders, H. (2021). Regional Innovation Scoreboard 2021. European Commission. https://doi.org/10.2873/674111

Horváth, D., & Szabó, R. Z. (2019). Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities? *Technological Forecasting and Social Change, 146*, 119–132. https://doi.org/10.1016/j.techfore.2019.05.021

Jondrow, J., Lovell, C. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics, 19*(2–3), 233–238. https://doi.org/10.1016/0304-4076(82)90004-5

Kleibrink, A., Niehaves, B., Palop, P., Sörvik, J., & Thapa, B. E. (2015). Regional ICT innovation in the European Union: Prioritization and performance (2008–2012). *Journal of the Knowledge Economy, 6*(2), 320–333. https://doi.org/10.1007/s13132-015-0240-0

Lehner, F., & Sundby, M.W. (2018). ICT Skills and Competencies for SMEs: Results from a Structured Literature Analysis on the Individual Level. In: Harteis, C. (eds) The Impact of Digitalization in the Workplace. Professional and Practice-based Learning, vol 21. Springer, Cham. https://doi.org/10.1007/978-3-319-63257-5_5

Li, L., Su, F., Zhang, W., & Mao, J. Y. (2018). Digital transformation by SME entrepreneurs: A capability perspective. *Information Systems Journal, 28*(6), 1129–1157. https://doi.org/10.1111/isj.12153

Martin, C., & Leurent, H. (2017). Technology and innovation for the future of production: Accelerating value creation. Geneva Switzerland. World Economic Forum. Available online at: http://www3.weforum.org/docs/WEF_White_Paper_Technology_Innovation_Future_of_Production_2017.pdf (Accessed 30/05/2022).

Martinez-Cillero, M., Lawless, M., O'Toole, C., & Slaymaker, R. (2020). Financial frictions and the SME investment gap: New survey evidence for Ireland. *Venture Capital, 22*(3), 239–259. https://doi.org/10.1080/13691066.2020.1771826

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 435–444,. https://doi.org/10.2307/2525757

Neokosmidis, I., Avaritsiotis, N., Ventoura, Z., & Varoutas, D. (2015). Assessment of the gap and (non-) Internet users evolution based on population biology dynamics. *Telecommunications Policy, 39*(1), 14–37. https://doi.org/10.1016/j.telpol.2014.10.006

Oberländer, M., Beinicke, A., & Bipp, T. (2020). Digital competencies: A review of the literature and applications in the workplace. *Computers & Education, 146*, 103752. https://doi.org/10.1016/j.compedu.2019.103752

Pellegrin, J., Colnot, L., Arendt, Ł., Bisaschi, L., Catalano, G., Martinaitis, Ž., & Micheletti, G. (2018). Research for REGI Committee-Digital Agenda and Cohesion Policy. European Parliament. Available online at: https://www.europarl.europa.eu/RegData/etudes/STUD/2018/617485/IPOL_STU(2018)617485_EN.pdf (Accessed 30/05/2022).

Reggi, L., & Scicchitano, S. (2014). Are EU regional digital strategies evidence-based? An analysis of the allocation of 2007–13 Structural Funds. *Telecommunications Policy, 38*(5–6), 530–538. https://doi.org/10.1016/j.telpol.2013.12.007

Reggi, L., & Gil-Garcia, J. R. (2021). Addressing territorial digital divides through ICT strategies: Are investment decisions consistent with local needs? *Government Information Quarterly, 38*(2), 101562. https://doi.org/10.1016/j.giq.2020.101562

Reis, J., Amorim, M., Melão, N., & Matos, P. (2018). Digital transformation: a literature review and guidelines for future research. In World conference on information systems and technologies (pp. 411–421). Springer, Cham. https://doi.org/10.1007/978-3-319-77703-0_41

Ruiz-Rodríguez, F., Lucendo-Monedero, A. L., & González-Relaño, R. (2018). Measurement and characterisation of the Digital Divide of Spanish regions at enterprise level. A comparative analysis with the European context. Telecommunications Policy, 42(3), 187–211.

Ulas, D. (2019). Digital transformation process and SMEs. *Procedia Computer Science, 158*, 662–671. https://doi.org/10.1016/j.procs.2019.09.101

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems, 28*(2), 118–144. https://doi.org/10.4324/9781003008637

Python Implementation of the Value-Based DEA Method

António Trigo, Maria Gouveia, and Carla Henriques

Abstract This paper is aimed at presenting the Python implementation of the Value-Based Data Envelopment Analysis (VBDEA) method, which was designed to evaluate the efficiency of decision-making units (DMUs). This methodological framework explores the links between data envelopment analysis (DEA) and multi-criteria decision analysis (MCDA) and proposes a new perspective on the use of the additive DEA model using concepts from the multi-attribute value theory (MAVT). One of the major strengths of VBDEA over typical DEA methodologies is that it offers information on the main reasons behind DMUs' (in)efficiency. Additionally, this approach allows straightforwardly ranking of efficient and inefficient DMUs, since it relies on a super-efficiency model. Because of the use of value functions, besides allowing the incorporation of the decision-maker (DM)'s preferences, this methodology easily handles negative or null data. In this context, we illustrate the Python implementation of the method by reproducing the main results obtained by (Gouveia et al., Or Spectrum 38:743–767, 2016), when these authors evaluated the performance of 12 health units in a Portuguese region incorporating management preferences given by real DMs.

Keywords Value-based DEA · Python · Efficiency

A. Trigo (✉) · M. Gouveia · C. Henriques
Polytechnic of Coimbra, Coimbra Business School | ISCAC, 3045-601 Coimbra, Portugal
e-mail: antonio.trigo@gmail.com

M. Gouveia · C. Henriques
DEEC, INESC Coimbra, Rua Sílvio Lima, Polo II, 3030-290 Coimbra, Portugal

Faculty of Economics, Univ Coimbra, CeBER, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

A. Trigo
ALGORITMI Research Center, University of Minho, 4800-058 Guimarães, Portugal

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_4

1 Introduction

DEA was originally developed by Charnes et al. (1978) and is a nonparametric approach based on linear programming to evaluate the performances of the DMUs (homogeneous units under evaluation) considering multiple inputs (resources) and multiple outputs (outcomes). The classical DEA models are usually (input or output) oriented models, which return radial efficiency measures. In this context, the CCR (Charnes et al., 1978) and the BCC (Banker et al., 1984) DEA approaches are both oriented and radial models. Besides, there are also other non-oriented models, such as the additive model (Charnes et al., 1985), which identify inefficient DMUs but do not return an efficiency score. Generally, it can be either assumed constant returns to scale (CRS) or variable returns to scale (VRS). While the CCR model only allows accounting for CRS, the BCC and additive models enable the consideration of VRS.

Since these DEA models allow computing the projections of inefficient DMUs onto the efficient frontier, depending on the scales used to measure the input and output factors, their efficiency measure is very pessimistic, because the L_1 distance is being maximized. Besides, particularly in the case of the additive model, the efficiency measure obtained does not have an intuitive interpretation. The VBDEA was developed by Gouveia et al. (2008) to overcome some disadvantages of the additive model, namely, the scaling problem. Therefore, this paper describes the Python of the VBDEA method. This paper is organized as follows. The next Section describes the VBDEA method. Section 3 explains the Python implementation. Finally, some conclusions are conveyed, and future work developments are unveiled.

2 The VBDEA Method

The present paper describes the Python implementation of the VBDEA method proposed by Gouveia et al. (2008). This method combines the use of DEA with MAVT (Keeney & Raiffa, 1993), in the field of MCDA, as a way of incorporating the preference information provided by DMs, converting the inputs and outputs (viewed as criteria of evaluation) into value scales. The additive value functions are used to aggregate the values associated with each criterion. This transformation makes it possible to overcome the problem of scales, as all criteria are translated into value units. Furthermore, the weights used in the aggregation gain a specific meaning, as they are the scale coefficients of the value functions and determine the projection direction. The weights are chosen to benefit each DMU as much as possible, in the optimistic spirit of the BCC models. Finally, the efficiency measure of each DMU will have an intuitive meaning: interpreted as the "min–max regret".

Consider n DMUs $\{DMU_j : j = 1, \ldots, n\}$ evaluated according to their performance in a set of q criteria, with $q = m + p$, with $x_{ij} (i = 1, \ldots, m)$ to be minimized, and $y_{rj}(r = 1, \ldots, p)$ to be maximized. The conversion consists of, using MAVT concepts, to build partial value functions $\{v_c(DMU_j), c = 1, \ldots, q, j = 1, \ldots, n\}$.

Each of them is defined within the interval [0, 1] assuming that for each criteria c the worst performance, p_{cj}, $j = 1, \ldots, n$, has the value 0 and the best performance, p_{cj}, $j = 1, \ldots, n$, has the value 1, causing the maximization of all criteria. Subsequently, the criteria are gathered into a global value function, $V(DMU_j) = \sum_{c=1}^{q} w_c v_c(DMU_j)$, where $w_c \geq 0, \forall c = 1,\ldots, q$ and $\sum_{c=1}^{q} w_c = 1$ (by convention). The weights $w_1, \ldots, w_q$ considered in the additive value function are the scale coefficients and are settled in a way that each alternative minimizes the value difference from the best alternative, according to the "min–max regret" rule (Bell, 1982).

The VBDEA method comprises two phases after all factors have been converted into a value scale.

Phase 1: Compute the efficiency measure, d_k^*, for each DMU_k ($k = 1, \ldots, n$), and the corresponding weighting vector $\boldsymbol{w}_k^*$ by solving problem (1).

$$
\begin{aligned}
&\min_{d_k, w} d_k \\
&s.t. \sum_{c=1}^{q} w_c v_c(DMU_j) - \sum_{c=1}^{q} w_c v_c(DMU_k) \leq d_k,\ j = 1, \ldots, n;\ j \neq k \\
&\sum_{c=1}^{q} w_c = 1 \\
&w_c \geq 0, \forall c = 1, \ldots, q.
\end{aligned}
\tag{1}
$$

It is worth noting that Gouveia et al. (2013) included the concept of superefficiency (Andersen & Petersen, 1993) in formulation (1) to accommodate the discrimination of efficient DMUs.

The optimal value of the objective function, d_k^*, , is the value difference to the best of all DMUs (note that the best DMU will also depend on $\boldsymbol{w}_k^*$), excluding itself from the reference set. If $d_k{}^*$ is negative, then the DMU_k under evaluation is efficient. In the end, it is possible to rank the efficient DMUs by considering that the more negative the value of d_k^*, the more efficient is DMU_k.

Phase 2: If $d_k^* \geq 0$, then solve the "weighted additive" problem (2), using the optimal weighting vector resulting from Phase 1, $\boldsymbol{w}_k^*$, and determine the corresponding projected point of the DMU_k under evaluation.

If d_k^* is non-negative, then DMU_k is inefficient and a projection target can be obtained through the following problem:

$$\min_{\lambda,s} z_k = -\sum_{c=1}^{q} w_c^* s_c$$
$$s.t. \sum_{j=1, j\neq k}^{n} \lambda_j v_c(DMU_j) - s_c = v_c(DMU_k), c = 1, \ldots, q$$
$$\sum_{j=1, j\neq k}^{n} \lambda_j = 1 \qquad (2)$$
$$\lambda_j, s_c \geq 0, j = 1, \ldots, k-1, k+1, \ldots, n; c = 1, \ldots, q$$

The group of efficient DMUs that defines a convex combination with $\lambda_j > 0$ ($j = 1, \ldots, k-1, k+1, \ldots, n$) is called the set of "benchmarks" of DMU_k. This convex combination leads to a point on the efficient frontier that is better than DMU_k by a difference of value of s_c (slack) in each criterion c.

2.1 Elicitation of Value Functions and Weight Restrictions

In the VBDEA method, the objective of converting the criteria into value scales (linear/nonlinear value functions) is to reflect the preferences of the DMs, considering the generalization of the DEA methodology presented by Cook and Zhu (2009) that incorporates piecewise linear functions of input and output factors.

To convert the criteria into value scales we established two limits, M_c^L and M_c^U, to consider an acceptable higher tolerance value (in this case, $\delta = 10\%$). We choose $M_c^L < min\left\{p_{cj}^L, j = 1, \ldots, n\right\}$ and $M_c^U > max\left\{p_{cj}^U, j = 1, \ldots, n\right\}$, for each $c = 1, \ldots, q$, to set the 0 and 1 levels on the value scale, according to the type of factor, input or output. After that, we compute value functions setting the values for each $DMU_j,\ j = 1, \ldots,$ n using:

$$v_c(DMU_j) = \begin{cases} \frac{p_{cj} - M_c^L}{M_c^U - M_c^L}, & if\ the\ factor\ c\ is\ an\ output \\ \frac{M_c^U - p_{cj}}{M_c^U - M_c^L}, & if\ the\ factor\ c\ is\ an\ input \end{cases}, j = 1, \ldots, \text{n};\ c = 1, \ldots, q \qquad (3)$$

To build the piecewise linear functions or non-linear value functions, we extract the difference in the DMU value that corresponds to decreases in inputs or increases in outputs, rather than the utility of having those inputs available or outputs produced. In this way, we do not speak of absolute values, but relative values.

The elicitation protocol can be based on comparing the value of increasing an output (or decreasing an input) from a to b versus increasing the same output (or decreasing the same input) from a' to b', all other performance levels being equal, and asking the DM to adjust one of these four numbers so that the value increase is approximately equal. This is always possible assuming the functions are continuous and monotonic.

The DM's answers to the questions about the value differences between the performance levels in each factor allow extracting the value functions, which can be a piecewise linear approximation. When the DM's responses can be fitted into predefined curves, we use other functions (like logarithmic, or exponential functions).

For a better understanding of the method and its implementation, we will follow the process with a replication of an illustrative example by Gouveia et al (2016). The purpose of the study carried out by Gouveia et al (2016) was to evaluate the efficiency of 12 primary health care units monitored by the "Group of Health Centres" in Portugal, with data from 2010. The perspective under consideration, designated as Model 2 in that study, uses as inputs (costs): total cost collected to the National Health Service (NHS) with complementary means of diagnosis and treatment (x_{CMDT}); total medicine costs collected to the NHS (x_{MED}); total cost of human resources (x_{HR}) and medical costs not collected to the NHS; clinical consumables and other costs (x_{OC}) and the only output is the number of medical consultations for registered patients (y_{CONS}).

In the literature, we can find several techniques to obtain information regarding the DM's preferences to construct value functions in agreement with his/her answers (Goodwin & Wright, 1998; von Winterfeldt & Edwards, 1986), but the questions must be structured for each specific context.

Table 1 summarizes the performance levels corresponding to values 0.25, 0.5, and 0.75 (resulting from this type of dialogue), such that an improvement from level 0 to level 0.25 corresponds to the same value as an improvement from level 0.25 to 0.5, etc. The summary of the performance levels elicited to construct the value functions for the output factor is depicted in Table 1, as an example.

For x_{CMDT}, x_{MED}, and x_{HR}, the value functions were obtained by fitting a logarithmic function to match as well as possible the answers of the DM.

In the VBDEA method, the DMU under evaluation is free to choose the scale coefficients (weights) of the marginal value functions aggregated with an additive MAVT model, to become the best DMU (if possible) or to minimize the difference of value to the best DMU, i.e., getting the best possible efficiency score considering only the marginal values of the inputs and outputs. However, some factors may be

Table 1 Summary of the performance corresponding to different (and equally spaced) value levels for the output factor

Value	Output
	y_{CONS}
0	4000
0.25	12,000
0.5	18,000
0.75	22,000
1	25,000

Source Authors' own elaboration

disregarded from the assessment, as DMUs may assign zero weight to some factors, incompatible with the DM's preferences. Thus, it is necessary to consider the weight constraints, as they may better reflect the organizational objectives and, therefore, guarantee significant results closer to what the DM considers to be the best practices.

There are several approaches to defining weight restrictions. In this context, specifying appropriate weight restrictions can be a very challenging task (Podinovski, 2004; Salo & Hämäläinen, 2001). In the Value-Based DEA method, the weights used in the aggregation are the scale coefficients of the value functions reflecting possible value trade-offs between different factors. Assigning values to the scale coefficients requires a series of judgments obtained from the DM. Direct classification techniques should be avoided, as the value of these coefficients does not reflect the DM's intuitive notion of the importance of each criterion. On the contrary, they are heavily dependent on the performances chosen to represent levels 0 and 1 on the value scale. In MCDA, several valid protocols are known to elicit weight restrictions derived from the DM's preferences (Goodwin & Wright, 1998; von Winterfeldt & Edwards, 1986). In this case, the swing technique is simple and clear for the DM. The swing method begins by constructing two extreme hypotheses, P0 and P1, with the first displaying the worst performance (having value 0) in all criteria scales and the second the corresponding best performance (having value 1). The preference elicitation protocol consists in querying the DM to look at the potential gains from moving from P0 to P1 in each criterion and then deciding which criterion he/she prefers to shift to hypothesis P1. Suppose that the transition from hypothesis P0 to hypothesis P1 in a specified criterion is worth 100 units on a hypothetical scale. Then, the DM is asked to give a value (<100) to the second criterion moved to P1, then to the third criterion, and so on, until the last criterion is moved to P1. The procedure used in the paper that we are using as a reference was to obtain, firstly, a ranking of weights and, secondly, to establish a limit for the ratio between the weights ranked first and last, to avoid null weights.

Considering W to be the set of weight vectors compatible with the elicited ranking and limit, it is necessary to include the weight restrictions in Phase 1 adding to formulation (1) the constraint $(\boldsymbol{w_1}, \ldots, \boldsymbol{w_q}) \in W$. With this change in Phase 1, a necessary change is mandatory in the formulation of the problem solved in Phase 2. This change allows slacks to have negative values; otherwise, it might not be possible to keep the optimal value difference d_k^* resulting from formulation (1) including the weight restrictions.

Weight restrictions were elicited by asking the DM to compare the "swings" from values 0 to 1 as depicted in Table 2.

Table 2 Extreme performances associated with value levels 0 and 1 (Model 2)

Value level	x_{CMDT}	x_{MED}	x_{HR}	x_{OC}	y_{CONS}
$u(.) = 0$	1,600,000	5,500,000	3,000,000	600,000	4,000
$u(.) = 1$	150,000	800,000	800,000	50,000	25,000

Source Authors' own elaboration

The DM was asked to consider a unit with the performance level 0 for all factors and the question was: “If you could improve one and only one factor in level 1, what would it be?”. The DM answer was: x_{MED}. This allows the inference that w_{MED} is the highest scaling coefficient. By repeating this question successively for the remaining factors, the ranking of the coefficients of scale obtained was: $w_{\mathrm{MED}} \geq w_{\mathrm{CMDT}} \geq w_{\mathrm{HR}} \geq w_{\mathrm{OC}} \geq w_{\mathrm{CONS}}$.

The answer to the question “What would be the lowest amount h that would allow a unit with 25,000 medical consultations for registered patients and total medicine costs collected to the NHS of 5.5 million euros to be considered as having more value than a unit with 4000 medical consultations for registered patients and total medicine costs collected to the NHS of h?” was $h = 2.5$ million euros. This answer is translated into: $w_{\mathrm{CONS}}\, v_{\mathrm{CONS}}(25{,}000) + w_{\mathrm{MED}}\, v_{\mathrm{MED}}(5{,}500{,}000) \geq w_{\mathrm{CONS}}\, v_{\mathrm{CONS}}(4000) + w_{\mathrm{MED}}\, v_{\mathrm{MED}}(h)$. Substituting h in the previous expression yields: $w_{\mathrm{MED}} \leq 2.47\, w_{\mathrm{CONS}}$.

3 Python Implementation

The Python implementation of the VBDEA method was done using the Python programming language (Python.org, 2022), and Jupyter Notebook (Jupyter.org, 2022) and the written code can also be executed from the console.

The VBDEA method has several steps for its execution, namely loading the model, converting the performances from the original scale to the value scales, the calculation of the first step of the method, the calculation of the second step of the method, and the conversion of the performances from the value scale to the original value scales, so the user understands the improvement proposals for the units classified as inefficient. The implementation of the different steps of the method is presented below in five sections.

The Jupyter notebook with Python implementation and the files of the Model 2 used for demonstration purposes are available in a Git Hub repository available at https://github.com/atrigo/vbDEA_notebook (Trigo et al., 2022).

3.1 Load the Model from a File

At this first step, the model to be executed is imported from a text file by a Python script. Figure 1 presents an example of such a file, relative to the case that we are using to demonstrate the python implementation (Gouveia et al., 2016). This file has the same name as the model which is “Model 2”.

Files containing models to be run by the application must have the same format as shown in Fig. 1, which consists of the following: a first line with the name of the model; then, lines that have the function type to be used for the conversion of the values to and from the value scale, which can be of three types: linear of multiple scales, exponential or logarithmic. Afterward, the values for matrix A are defined;

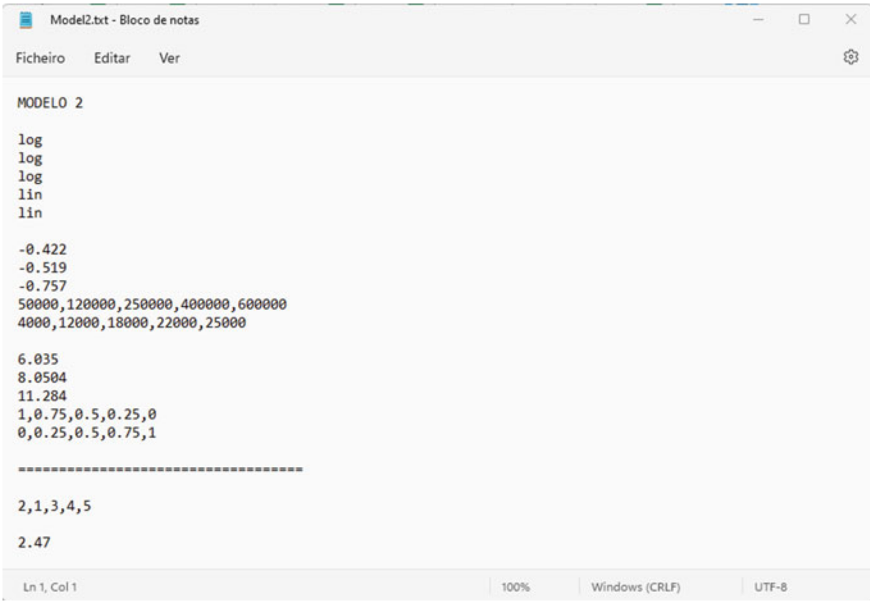


Fig. 1 Example of a text file containing the model. *Source* Authors' own elaboration

the next lines contain the values for matrix B, followed by a dashed line; and, finally, two lines with optional parameters: the first line with the importance of the factors, ordered from the most important (in the case of Fig. 1, the second factor) to the least important (in the case of Fig. 1, the last factor), and, the second line, with an optional parameter (in the case of Fig. 1, the value 2.47, which represents the limit for the ratio between the weights ranked first and last). Note that between all the model parameter definitions there are blank lines that must be respected for the Python script to work.

Figure 2 shows the output of the Python script after running this first step. If everything goes well, the output parameters defined in the text file can be seen.

3.2 Conversion of the Values to Value Scales

Once the code that loads the model to be executed has been created, the second step is the conversion of the performances in original scale values into value scales according to the chosen model functions.

Figure 3 depicts the file with the DMUs' performances in original scales. Figure 4 shows the file with the conversion of performances from original scales into value scales with the algorithm created in the Python language. The Python algorithm reads a file with the name <[modelname]_originals.csv> and returns a file with the name <[modelname]_valuescale.csv> with the values converted.

```
✓ 0.1s                                                    Python
Model: Model2
Function list:
['log', 'log', 'log', 'lin', 'lin']
A matrice:
[-0.422, -0.519, -0.757, [50000.0, 120000.0, 250000.0, 400000.0, 600000.0], [4000.0,
12000.0, 18000.0, 22000.0, 25000.0]]
B matrice:
[6.035, 8.0504, 11.284, [1.0, 0.75, 0.5, 0.25, 0.0], [0.0, 0.25, 0.5, 0.75, 1.0]]
Priority factor array (most important to least important):
[2, 1, 3, 4, 5]
Limit for the ratio between the weights ranked first and last:
2.47
```

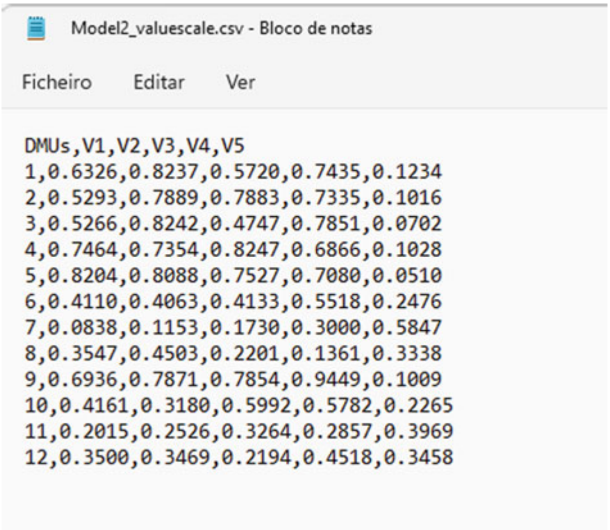
Fig. 2 Python output from file loading. *Source* Authors' own elaboration

```
Model2_Originals.csv - Bloco de notas
Ficheiro   Editar   Ver

DMUs,X1,X2,X3,X4,Y1
1,362900,1114935,1398033,123382,8962
2,463549,1192330,1050558,128606,8439
3,466561,1113734,1589818,101750,6247
4,277140,1321779,1001297,152974,7291
5,232565,1147295,1101200,141864,5631
6,613535,2491923,1724028,218890,11942
7,1332317,4365605,2368401,360011,19356
8,701121,2289334,2225321,491140,14011
9,314090,1196353,1054564,65416,7230
10,606148,2954054,1348651,203098,11436
11,1007909,3350321,1933791,371468,16351
12,708936,2793686,2227392,278936,14300
```

Fig. 3 Input file with the performances in original scales. *Source* Authors' own elaboration

To better understand the values of the above files, Table 3 is presented, which has the original factors' names, descriptions, and types, and corresponding codification based on the files depicted in Figs. 3 and 4.



Model2_valuescale.csv - Bloco de notas

Ficheiro Editar Ver

```
DMUs,V1,V2,V3,V4,V5
1,0.6326,0.8237,0.5720,0.7435,0.1234
2,0.5293,0.7889,0.7883,0.7335,0.1016
3,0.5266,0.8242,0.4747,0.7851,0.0702
4,0.7464,0.7354,0.8247,0.6866,0.1028
5,0.8204,0.8088,0.7527,0.7080,0.0510
6,0.4110,0.4063,0.4133,0.5518,0.2476
7,0.0838,0.1153,0.1730,0.3000,0.5847
8,0.3547,0.4503,0.2201,0.1361,0.3338
9,0.6936,0.7871,0.7854,0.9449,0.1009
10,0.4161,0.3180,0.5992,0.5782,0.2265
11,0.2015,0.2526,0.3264,0.2857,0.3969
12,0.3500,0.3469,0.2194,0.4518,0.3458
```

Fig. 4 Output file with the performances converted into value scales. *Source* Authors' own elaboration

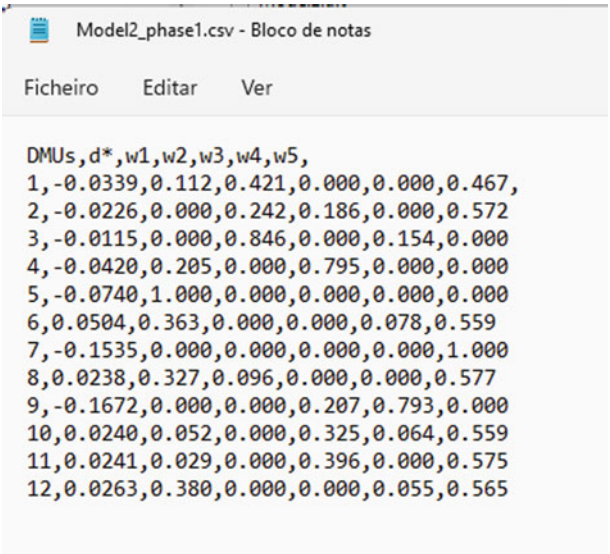
Table 3 Enumeration of Model 2 factors

Factors	Factors description	Variable type	Variable original codification	Variable codification in the value scale
x_{CMDT}	Total cost billed to the National Health Service (NHS)	Input	*X1*	*V1*
x_{MED}	Total medicine costs billed to the NHS	Input	*X2*	*V2*
x_{HR}	Total cost of human resources	Input	*X3*	*V3*
x_{OC}	Medicine costs not billed to the NHS, clinical consumables and other costs	Input	*X4*	*V4*
y_{CONS}	Number of medical consultations for registered patients	Output	*Y1*	*V5*

Source Authors' own elaboration

3.3 Calculations of the First Step of the Method

After converting the values into value scales, we are ready to run the first phase of the model which consists of computing the efficiency measure, d_k^*, for each DMU_k


Model2_phase1.csv - Bloco de notas

Ficheiro Editar Ver

```
DMUs,d*,w1,w2,w3,w4,w5,
1,-0.0339,0.112,0.421,0.000,0.000,0.467,
2,-0.0226,0.000,0.242,0.186,0.000,0.572
3,-0.0115,0.000,0.846,0.000,0.154,0.000
4,-0.0420,0.205,0.000,0.795,0.000,0.000
5,-0.0740,1.000,0.000,0.000,0.000,0.000
6,0.0504,0.363,0.000,0.000,0.078,0.559
7,-0.1535,0.000,0.000,0.000,0.000,1.000
8,0.0238,0.327,0.096,0.000,0.000,0.577
9,-0.1672,0.000,0.000,0.207,0.793,0.000
10,0.0240,0.052,0.000,0.325,0.064,0.559
11,0.0241,0.029,0.000,0.396,0.000,0.575
12,0.0263,0.380,0.000,0.000,0.055,0.565
```

Fig. 5 Output file from phase 1. *Source* Authors' own elaboration

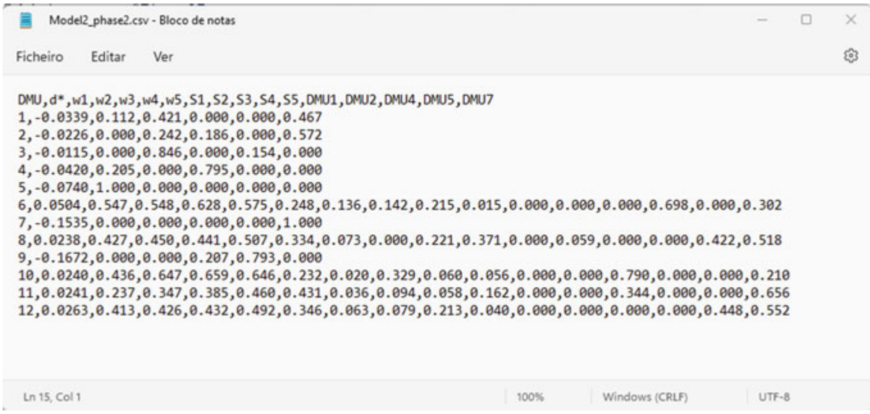
($k = 1, \ldots, n$), and the corresponding weighting vector w_k^* by solving linear problem (1), as previously described in Sect. 2.

The implementation of this part of the algorithm had to use a Python solver, in this case, the *linprog* function of the package *scipy.optimize* (The SciPy community, 2022[1]), which has several functions for optimization, in addition to using the numpy and pandas' packages already used before in the previous sections.

The results from the first phase of the model are presented in Fig. 5 and the ranking of units is: DMU 9 $\succ$ DMU 7 $\succ$ DMU 5 $\succ$ DMU 4 $\succ$ DMU 1 $\succ$ DMU 2 $\succ$ DMU 3 $\succ$ DMU 8 $\succ$ DMU 10 $\succ$ DMU 11 $\succ$ DMU 12 $\succ$ DMU 6, where the first seven DMUs are efficient, because they have $d^* < 0$. The lower the value of d^* the better, and if d^* is negative, then the DMU under analysis is efficient; otherwise, it is inefficient.

The DMUs freely choose their weights to become the best DMU (if possible) or to minimize the difference in value for the best DMU. There are units that disregard some factors from evaluation, such as DMU 5 and DMU 7, that considered only one of the five factors to be ranked as efficient, namely $w^*_{CMDT} = 1$ and $w^*_{CONS} = 1$, respectively.

[1] https://docs.scipy.org/doc/scipy/reference/optimize.html.



Model2_phase2.csv - Bloco de notas

Ficheiro Editar Ver

```
DMU,d*,w1,w2,w3,w4,w5,S1,S2,S3,S4,S5,DMU1,DMU2,DMU4,DMU5,DMU7
1,-0.0339,0.112,0.421,0.000,0.000,0.467
2,-0.0226,0.000,0.242,0.186,0.000,0.572
3,-0.0115,0.000,0.846,0.000,0.154,0.000
4,-0.0420,0.205,0.000,0.795,0.000,0.000
5,-0.0740,1.000,0.000,0.000,0.000,0.000
6,0.0504,0.547,0.548,0.628,0.575,0.248,0.136,0.142,0.215,0.015,0.000,0.000,0.000,0.698,0.000,0.302
7,-0.1535,0.000,0.000,0.000,0.000,1.000
8,0.0238,0.427,0.450,0.441,0.507,0.334,0.073,0.000,0.221,0.371,0.000,0.059,0.000,0.000,0.422,0.518
9,-0.1672,0.000,0.000,0.207,0.793,0.000
10,0.0240,0.436,0.647,0.659,0.646,0.232,0.020,0.329,0.060,0.056,0.000,0.000,0.790,0.000,0.000,0.210
11,0.0241,0.237,0.347,0.385,0.460,0.431,0.036,0.094,0.058,0.162,0.000,0.000,0.344,0.000,0.000,0.656
12,0.0263,0.413,0.426,0.432,0.492,0.346,0.063,0.079,0.213,0.040,0.000,0.000,0.000,0.000,0.448,0.552
```

Ln 15, Col 1 100% Windows (CRLF) UTF-8

Fig. 6 Output file from phase 2. *Source* Authors' own elaboration

3.4 Calculations of the Second Stage of the Method

In Phase 2 of the VBDEA method, the optimal weighting vector is used to solve the problem with formulation (2) for the DMUs classified as inefficient. The solution is a proposed efficiency target (projection) for each inefficient DMU. To achieve the efficient status, these inefficient DMUs must change their value in each factor by the value indicated by s*.

In our example, the DMU that is most often selected as a benchmark is DMU 7, and, for example, DMU 6 is inefficient, and it is projected onto the efficiency frontier in a target obtained by a linear combination of DMUs 4 and 7.

Figure 6 shows the outputs of the second phase of the model. In this file, for model readability reasons, the weights (w1, w2, w3, w4, and w5) calculated in the first stage of the model are also visible. Table 4 shows the same output but formatted in a tabular form.

3.5 Conversion of Values into the Original Scale

Table 5 depicts the values of the slacks in their original value scales.

As the slacks present positive values only for inputs, these values must be subtracted from the values of the performances in the original scale, thus obtaining the projected points in the efficiency frontier. The slack values translate the reductions to be implemented in the inputs in the sense that each of the inefficient DMUs manages to be at the level of those that are operating efficiently, i.e., those that are examples of best practices.

Table 4 Tabular output of Phase 2

DMU	d*	w1	w2	w3	w4	w5	S1	S2	S3	S4	S5	DMU1	DMU2	DMU4	DMU5	DMU7
1	−0.034	0.11	0.42	0	0	0.47										
2	−0.023	0	0.24	0.19	0	0.57										
3	−0.012	0	0.85	0	0.15	0										
4	−0.042	0.21	0	0.8	0	0										
5	−0.074	1	0	0	0	0										
6	0.0504	0.55	0.55	0.63	0.58	0.25	0.14	0.14	0.22	0.02	0	0	0	0.698	0	0.302
7	−0.154	0	0	0	0	1										
8	0.0238	0.43	0.45	0.44	0.51	0.33	0.07	0	0.22	0.37	0	0.059	0	0	0.422	0.518
9	−0.167	0	0	0.21	0.79	0										
10	0.024	0.44	0.65	0.66	0.65	0.23	0.02	0.33	0.06	0.06	0	0	0.79	0	0	0.21
11	0.0241	0.24	0.35	0.39	0.46	0.43	0.04	0.09	0.06	0.16	0	0	0.344	0	0	0.656
12	0.0263	0.41	0.43	0.43	0.49	0.35	0.06	0.08	0.21	0.04	0	0	0	0	0.448	0.552
Number of times of benchmark												**1**	**2**	**1**	**2**	**5**

Source Authors' own elaboration

Table 5 Slacks in their original value scales

DMU	S1	S2	S3	S4	S5
1					
2					
3					
4					
5					
6	83,520	1,160,608	3,602,396	588,000	0
7					
8	6336	1,624,885	3,560,989	153,200	0
9					
10	4640	745,139	4,856,172	555,200	0
11	5152	1,300,421	4,874,921	470,400	0
12	6016	1,347,476	3,616,305	568,000	0

Source Authors' own elaboration

4 Conclusions and Further Research

The purpose of this work is to describe the Python implementation of the VBDEA approach. This methodological framework explores the connections between DEA and MCDA and presents a fresh viewpoint on the application of the additive DEA model based on MAVT. One of the major advantages of VBDEA over traditional DEA approaches is that it provides information on the leading causes of DMUs' (in)efficiency. Furthermore, because it is based on a super-efficiency model, this technique has a higher discriminatory power since it allows ranking both efficient and inefficient DMUs. With the use of value functions, this technique, in addition to permitting the inclusion of the DMs' preferences, readily handles negative or null data. In this regard, we display the Python implementation of the approach by replicating the major findings of Gouveia et al. (2016), who assessed the efficiency of 12 health facilities in a Portuguese region using management preferences provided by real DMs.

Future work is currently underway to further develop the algorithm presented to make it freely available in a web application so that it can be used by different types of users (https://adept.iscac.pt).

Acknowledgements This work has been funded by European Regional Development Fund in the framework of Portugal 2020—Programa Operacional Assistência Técnica (POAT 2020), under project POAT-01-6177-FEDER-000044 ADEPT: Avaliação de Políticas de Intervenção Co-financiadas em Empresas. INESC Coimbra and CeBER are supported by the Portuguese Foundation for Science and Technology funds through Projects UID/MULTI/00308/2020 and UIDB/05037/2020, respectively.

References

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science, 39*(10), 1261–1264.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science, 30*, 1078–1092.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research, 30*(2), 961–981.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics, 30*(1–2), 91–107. https://doi.org/10.1016/0304-4076(85)90133-2

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*, 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Cook, W. D., & Zhu, J. (2009). Piecewise linear output measures in DEA (third revision). *European Journal of Operational Research, 197*, 312–319.

Goodwin, P, & Wright, G. (1998). *Decision analysis for management judgment* (2nd edn.). Chichester: Wiley.

Gouveia, M. C., Dias, L. C., & Antunes, C. H. (2008). Additive DEA based on MCDA with imprecise information. *Journal of the Operational Research Society, 59*(1), 54–63. https://doi.org/10.1057/palgrave.jors.2602317

Gouveia, M. C., Dias, L. C., & Antunes, C. H. (2013). Super-efficiency and stability intervals in additive DEA. *Journal of the Operational Research Society, 64*(1), 86–96. https://doi.org/10.1057/jors.2012.19

Gouveia, M. C., Dias, L. C., Antunes, C. H., Mota, M. A., Duarte, E. M., & Tenreiro, E. M. (2016). An application of value-based DEA to identify the best practices in primary health care. *Or Spectrum, 38*(3), 743–767.

Jupyter Notebook. (2022). Jupyter.org. Retrieved July 28, 2022, from https://www.jupyter.org.

Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge Univer- sity Press.

Podinovski, V. V. (2004). Production trade-offs and weight restrictions in data envelopment analysis. *Journal of the Operational Research Society, 55*, 1311–1322.

Python Programming Language. (2022). Python.Org. Retrieved July 28, 2022, from https://www.python.org.

Salo, A. A., & Hämäläinen, R. P. (2001). Preference Ratios in Multiattribute Evaluation (PRIME)—elicitation and decision procedures under incomplete information. *IEEE Transactions on Systems Man and Cybernetics Part A, 31*, 533–545.

The SciPy community. (2022). Optimization and root finding (scipy.optimize)#. Optimization and root finding (scipy.optimize)—SciPy v1.9.0 Manual. Retrieved July 29, 2022, from https://docs.scipy.org/doc/scipy/reference/optimize.html.

Trigo, A., Henriques C., & Gouveia, M. (2022). vbDEA notebook. Retrieved July 28, 2022, from https://github.com/atrigo/vbDEA_notebook.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis behavioral research*. Cambridge University Press.

Are ERDF Efficient in Strengthening the Switch to a Low-Carbon Economy? Some Insights with Value-Based Data Envelopment Analysis

Maria Gouveia, Carla Henriques, and Ana Amaro

Abstract We assessed the execution of European Regional Development Funds (ERDF) allocated to promote a Low-carbon economy (LCE) in 23 EU Member States (MS). Each MS is evaluated using the Value-Based Data Envelopment Analysis (VBDEA) method. In the first stage, the efficient MS were identified, and the major reasons that might affect the efficient performance of the ERDF. From the results obtained, 43% of the MS were deemed efficient in the application of ERDF committed to fostering an LCE, and these results were mostly justified by their financial spending rate. At the second stage of the analysis, the changes that needed to be done by inefficient MS to "try and replicate" their efficient counterparts were computed. Furthermore, from the robustness assessment conducted it was possible to show that with thresholds of $\delta = 5\%$ and $\delta = 10\%$, 22% of the MS managed to attain a robust efficiency. While Spain is the leading country in terms of robustness efficiency, Romania (robustly inefficient for $\delta = 5\%$), Hungary, and the Czech Republic (the worst in the ranking of inefficient MS) could not apply these funds properly. Given this information, the EU should continue to push policies that secure financial opportunities from engaging in LCE, particularly for MS with limited financial capacities, while still supplying them with improved funding mechanisms and technical expertise.

Keywords Low-carbon economy · ERDF · Value-based DEA · Robustness analysis

M. Gouveia (✉) · C. Henriques · A. Amaro
Polytechnic of Coimbra, Coimbra Business School|ISCAC, 3045-601 Coimbra, Portugal
e-mail: mgouveia@iscac.pt

C. Henriques
e-mail: chenriques@iscac.pt

M. Gouveia · C. Henriques
INESC Coimbra, DEEC, Rua Sílvio Lima, Polo II, 3030-290 Coimbra, Portugal

Univ Coimbra, CeBER, Faculty of Economics, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

A. Amaro
CEGIST-IST-UL, 1049-001 Lisboa, Portugal

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_5

1 Introduction

The EU policy for mitigating the effects of climate change needs to urgently shift towards an LCE. Simply explained, an LCE is an economy whose organization is sustained by activities that emit minimal quantities of CO_2 into the environment (Levy, 2010). The cohesion policy has been supporting the shift to an LCE, but in the 2014–2020 period this support has expanded dramatically, owing in part to the granting of specific funding for this purpose (Henriques et al., 2022a). As a result, assessment plays an important role in cohesion policy-making since it supports policy planning and development while also delivering solid data about the outcomes and effects of the projects undertaken. In this framework, the Data Envelopment Analysis (DEA) method has been particularly useful in the assessment of the OPs devoted to the "competitiveness of Small and Medium Sized-Enterprises (SMEs)" (Gouveia et al., 2021) to an LCE in SMEs (Henriques et al., 2022a) and to research and in Innovation in SMEs (Henriques et al., 2022b). When conducting an efficiency assessment through the DEA approach, management authorities (MA) will be able to pinpoint the OPs viewed as a reference of best practices and the required changes that have to take place for the set of indicators, which will allow transforming inefficient OPs into efficient ones across the programmatic horizon. Additionally, DEA can likewise be used in the efficiency evaluation of the LCE across distinctive settings, specifically at the national (Chen et al., 2020; Liu & Liu, 2016; Zhang et al., 2019, 2020), regional (Meng et al., 2018; Zhang et al., 2017), and sectoral (Zha, et al., 2019) levels. In this regard, hardly any of the papers available in the scientific literature use DEA in the evaluation of the MS global efficiency in the use of ERDF to promote an LCE. Furthermore, as far as we are aware, the VBDEA has never been employed in this situation. One of the major advantages of the VBDEA above traditional DEA techniques is that it provides an additional understanding of the underlying causes of (in)efficiency. This technique also enables tackling negative or null data, studying the robustness of the results, and incorporating the DM's preferences in the appraisal through the use of value functions. As a result, we want to contribute to the literature by undertaking an efficiency evaluation of the application of ERDF committed to LCE over 23 European countries. In summary, our main research questions are given below:

RQ1: "What are the main reasons for the (in)efficiency in the utilization of ERDF granted to promote an LCE in EU countries?"

RQ2: "Which countries were considered benchmarks during the last programmatic timeframe?"

RQ3: "Which MS performs better in terms of robustness?"

This paper's structure is as follows. Section 2 explains the basic premises underlying the methodologies offered to appraise the implementation of the ERDF in the countries under scrutiny. Section 3 explains why the criteria employed herein were chosen. Section 4 highlights the main results. Section 5 reports the main conclusions, discusses prospective political repercussions, highlights important flaws, and proposes future study topics.

2 Methodology

We employ a DEA-based approach in this work, which is a classical optimization tool that generates an efficiency frontier by evaluating homogenous decision-making units (DMUs), in this case, the MS. This method enables the consideration of numerous criteria (to be maximized or minimized). This type of tool can provide relevant information, such as the reasons behind (in)efficiency, the efficient peers of inefficient DMUs, and the needed modifications to the criteria used in the assessment to reach efficiency. We focus on the Gouveia et al. (2008) VBDEA model, which integrates the application of DEA with multi-attribute value theory (MAVT) (Keeney & Raiffa, 1993). The VBDEA model, proposed by Gouveia et al. (2008), addresses the scales challenge and the lack of understanding of the value produced by the weighted additive model (Ali et al., 1995). In the realm of MCDA, this last technique allows for the incorporation of the DM's preferences by turning the criteria into value scales. This transformation is very useful for dealing with negative or null data. Besides, inspired by the concept of superefficiency in DEA models (Gouveia et al., 2013), the VBDEA approach allows ranking, in a single step, all DMUs, even the efficient ones, and enables to consider the robustness analysis of the (in)efficiency values obtained.

The VBDEA method involves two stages after all factors have been transformed into value scales. At the first stage, the optimal value difference to the best of all DMUs, excluding itself is computed, i.e., the efficiency score is obtained. If this distance is negative, then the DMU under scrutiny is efficient; otherwise, it is inefficient. The ranking of the DMUs can then be done from the most efficient to the less efficient from the most negative values to the most positive values attained for this value difference. In the second stage, the reference set of efficient DMUs is computed for each inefficient DMU, by instantiating this second model with the optimal values obtained previously. Further details on this method and the corresponding software might be found in Chap. "Python Implementation of the Value-Based DEA Method" of this book.

3 Data

The criteria used in this study were suggested by Henriques et al. (2022a) at the OP level. The numbers evaluated are total figures at the MS level spanning various years reported on November 19, 2021 (corresponding to the programming period of 2014–2020), because they are the most up-to-date statistics for the accomplishment criteria. Just those MS with comprehensive data on ERDF grants were examined in the research. The criteria chosen for evaluating the efficiency of fund execution were drawn from a set of shared criteria officially mandated by the EU (European Commission, 2014) and are explained in Table 1.

Data on the descriptive statistics of these factors are given in Table 2.

Table 1 Criteria used

	EU co-financing	Total eligible spending	Eligible cost decided	GHG reduction
Description	Percentage of EU financing (calculated as an average)	Eligible costs validated	Financial resources assigned	Estimated annual decrease of GHG
Type of factor	To minimize	To maximize	To minimize	To maximize
Unit	%	Euro	Euro	Tons of CO_2 equivalent
Source	(a)	(a)	(a)	(b), (c)
Explanation	EU subsidy dependency	OPs' execution	OPs' execution	Reaching an LCE

Source Authors' own elaboration based on Henriques et al. (2022a)
(a) List of Structural Funds financial implemented data. Available at: https://cohesiondata.ec.europa.eu/2014-2020/ESIF-2014-2020-Finance-Implementation-Details/99js-gm52
(b) List of common indicators legally required and listed in the annexes to the ERDF, Cohesion Fund and ETC regulations. Available at: https://ec.europa.eu/regional_policy/sources/docoffic/2014/working/wd_2014_en.pdf
(c) List of Structural Funds achievement data. Available at: https://cohesiondata.ec.europa.eu/2014-2020/ESIF-2014-2020-Achievement-Details/aesb-873i

Table 2 Descriptive statistics of data in their original performances

Statistics	EU co-financing	Eligible cost decided	Total eligible spending	GHG reduction
Mean	64.65	5,003,186,122.35	1,744,268,042.09	181,264.20
Median	65.56	1,740,681,178.00	571,461,318.00	82,646.53
Standard deviation	17.09	5,732,200,879.94	2,096,538,613.08	215,620.69
Minimum	29.46	99,506,488.00	28,265,375.00	37.16
Maximum	85.00	17,105,000,000.00	7,016,881,169.00	829,915.74
Count	23	23	23	23

Source Authors' own elaboration

4 Discussion of Results

We could have used non-linear value functions, but because we didn't have an actual DM, we decided to convert all of the criteria into linear value functions, showing neutral preferences. This conversion considered two bounds, M_c^L and M_c^U. The lower and upper bounds were obtained, respectively, as $M_c^L < min\left\{p_{cj}^L, j = 1, \ldots, n\right\}$ and $M_c^U > max\left\{p_{cj}^U, j = 1, \ldots, n\right\}$, where $p_{cj}^L = p_{cj}(1-\delta) \leq p_{cj} \leq p_{cj}(1+\delta) = p_{cj}^U$, with $\delta = 10\%$ and p_{cj} is the performance of criterion c for DMU_j, $c = 1, \ldots, q$ and $j = 1, \ldots, n$. The value scales are then set between the 0 and

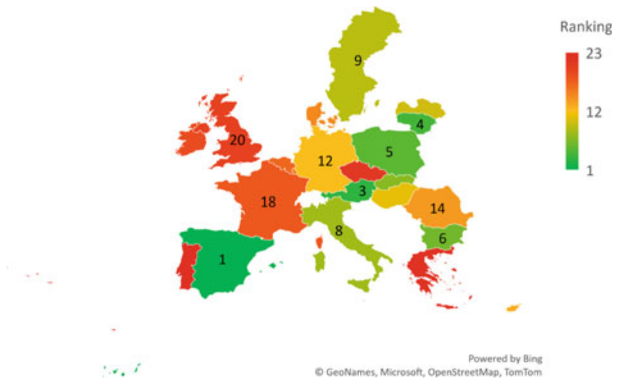


Fig. 1 Ranking of the countries according to in(efficiency) scores. *Source* Authors' own elaboration

1 levels depending on the type of criteria (maximized or minimized)—see expression (1). Afterward, we obtain the values for each DMU_j, $j = 1, \ldots, n$ employing (1):

$$v_c(DMU_j) = \begin{cases} \dfrac{pcj - M_c^L}{M_c^U - M_c^L}, & \text{if } c \text{ is being maximized} \\ \dfrac{M_c^U - pcj}{M_c^U - M_c^L}, & \text{if } c \text{ is being minimized} \end{cases}, j = 1, \ldots, n;\ c = 1, \ldots, q \quad (1)$$

Subsequently, the VBDEA is instantiated with the performance values thus obtained leading to the computation of the results depicted in Fig. 1.

From Fig. 1 it is possible to conclude that only 10 countries attain an efficient status, corresponding to Spain, Malta, Austria, Lithuania, Polonia, Bulgaria, Slovakia, Italy, Sweden, and Latvia (countries identified with different shades of green). The countries that show the best performance are Spain, followed by Malta, and Austria. These MS have outstanding efficiency levels. Besides, as it will be seen further these are also the most robust from the set of efficient MS.

Figure 1 depicts the outcomes as well for inefficient MS. In this case, Portugal, Greece, and the Czech Republic show the worst performance.

Phase 1 of VBDEA allows obtaining the efficiency scores that enable ranking both efficient and inefficient MS, as well as the corresponding weighting vectors that reflect the importance given to each criterion to attain the best efficiency score possible—Fig. 2.

The indicator highly sought by MS to attain the greatest efficient performance was "Eligible costs decided" (w2) followed by "Total eligible spending" (w3) (Fig. 2). Nevertheless, neither of the three leading scoring nations elected the criterion "Eligible cost decided"—see Figs. 2 and 3b. Spain only prioritized the criteria being maximized ("Total eligible spending" and "GHG emission reduction") (w3 = 0.363 and w4 = 0.637). Malta ranked 3rd based solely on "Total eligible spending" (w3 = 1)—see Fig. 3c, whereas Austria ranked 3rd based solely on "EU co-financing" (w1 = 1) (see Fig. 3a). Only four of the ten efficient MS (Bulgaria, Spain, Italy, and

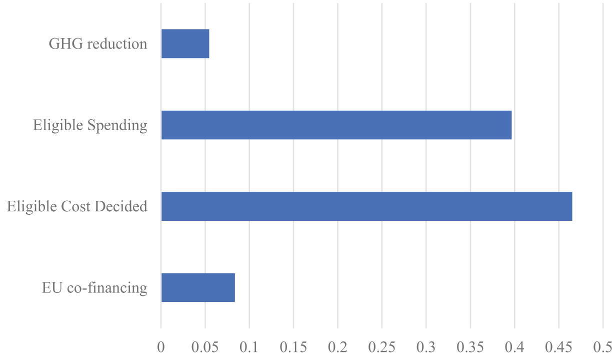
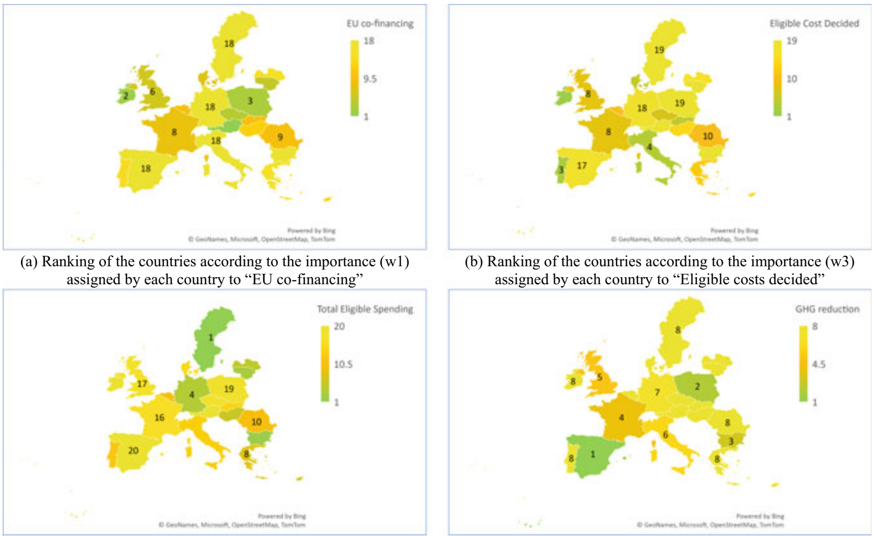


Fig. 2 Average weights obtained with VBDEA. *Source* Authors' own elaboration

Poland) choose "GHG reduction" as a crucial factor for achieving efficiency—see Fig. 3d.

The top three MS usually chosen as benchmarks for the non-efficient MS are Slovakia (8 times), Malta (5 times), and Austria (5 times)—see Fig. 4. Furthermore, two of the four Visegrad MS are efficient in the execution of ERDF funding committed to an LCE, with just one being more regularly designated as a benchmark (Slovakia)—see Fig. 4.



(a) Ranking of the countries according to the importance (w1) assigned by each country to "EU co-financing"

(b) Ranking of the countries according to the importance (w3) assigned by each country to "Eligible costs decided"

(c) Ranking of the countries according to the importance (w2) assigned by each country to "Eligible spending"

(d) Ranking of the countries according to the importance (w4) assigned by each country to "GHG reduction"

Fig. 3 Results obtained according to the weight vectors computed with VBDEA (ranking in decreasing order). *Source* Authors' own elaboration

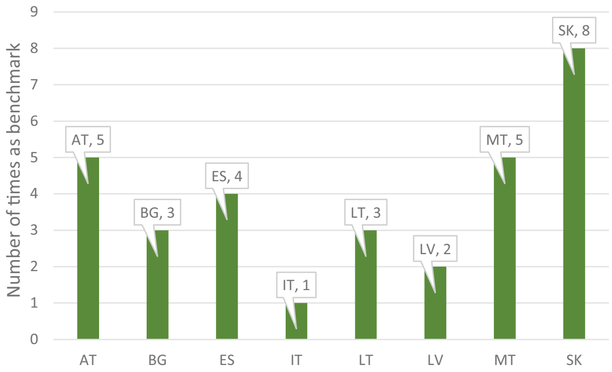


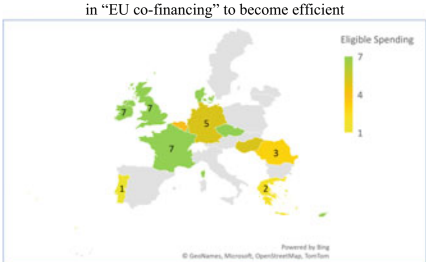
Fig. 4 Number of times each MS has been selected as benchmark. *Source* Authors' own elaboration

In Phase 2, the efficient targets (projections) are computed for each inefficient MS. To reach efficiency, these inefficient MS must adjust their performance values in each criterion according to the value obtained for the slacks—see Fig. 5.

The Czech Republic is the country that needs to make the biggest GHG reduction of all countries in the sample, followed by Ireland and Luxembourg (Fig. 5d). Additionally, the only inefficient MS that do not need to further reduce GHG emissions are Germany, France, and the UK. These MS inefficiency resides in their overallocation

(a) Ranking of the countries according to the required change (s1) in "EU co-financing" to become efficient

(b) Ranking of the countries according to the required change (s2) in "Eligible costs decided" to become efficient



(c) Ranking of the countries according to the required change (s3) in "Eligible spending"to become efficient

(d) Ranking of the countries according to the required change (s4) in "GHG reduction" to become efficient

Fig. 5 Results obtained according to the slacks computed for inefficient countries with VBDEA (ranking in decreasing order). *Source* Authors' own elaboration

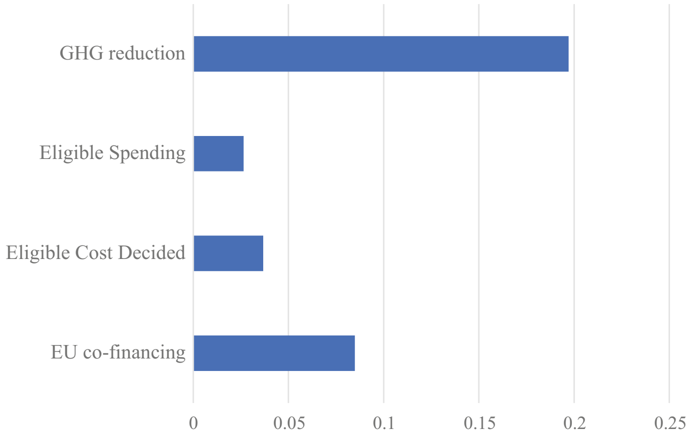


Fig. 6 Average required adjustments given by the slacks. *Source* Authors' own elaboration

of EU funding committed to fostering an LCE (Fig. 5a, b). Germany and France, for example, see a robust climate policy agenda as beneficial to their national economy (Bąk et al., 2021).

From Fig. 6 it can be established that the major required average changes are "GHG reduction", followed by the dependence on "EU co-financing".

In Table 3 we can see the improvements to be made in the original performance scale and the projections on the efficient frontier of all the inefficient countries.

The values considered to represent the criteria are occasionally uncertain. In such cases, the original DEA model is transformed into two models, thus enabling to obtain the upper and lower bounds of the efficiency scores. In the first model (the worst scenario), all DMUs' criteria being maximized are raised while all the criteria being minimized are lowered, except for the DMU under study (i.e., DMU_k worsens its efficiency performance while the remaining DMUs improve their efficiency performance). In the best scenario, the opposite case is considered (Gouveia et al., 2013). The robustness assessment of the efficiency scores for each MS is illustrated in Fig. 7, applying a rate of change of $\delta = 5\%$ and $\delta = 10\%$.

Spain, Austria, Malta, Lithuania, Bulgaria, and Sweden are robustly efficient (in decreasing order) for both data perturbations used (5 and 10%). Portugal, Belgium, Luxembourg, Greece, and Cyprus are robustly inefficient (in decreasing order) for both tolerances. Latvia is only robustly efficient for a tolerance of 5%, but just potentially efficient with a tolerance of 10% applied to all the criteria. Germany and Romania are robustly inefficient for a data perturbation of 5% and potentially efficient for a data perturbation of 10%. The remaining countries are potentially efficient for both data perturbations. Furthermore, this type of analysis allows concluding that Spain is by far the most robust MS in terms of efficiency. Curiously, Slovakia, which was most frequently selected as a benchmark (see Fig. 4) is just potentially efficient for all data perturbations.

Table 3 Adjustments in original performances and targets on the efficient frontier

Slacks in original performances					Projections			
DMU	s1	s2	s3	s4	EU co-financing	Eligible cost decided	Total eligible spending	GHG reduction
BE	−10.54	−612,108,067.00	96,390,688.00	138,879.69	29.46	1,128,573,111.00	417,572,434.00	138,916.85
CY	0.00	−243,998,058.00	0.00	64,532.92	67.50	392,204,743.00	186,165,272.00	72,775.88
CZ	−12.32	−627,994,096.00	0.00	713,452.80	45.05	13,745,000,000.00	4,962,960,249.00	796,099.33
DE	−8.83	−70,717,200.00	8,982,977.08	0.00	46.17	193,477,203.00	64,476,257.80	250,200.12
DK	−5.88	−398,001,267.00	0.00	36,110.55	59.68	9,752,309,210.00	4,032,596,628.00	87,115.55
FR	0.00	−451,573,188.00	0.00	0.00	76.25	6,444,850,860.00	2,409,491,434.00	349,456.89
GR	0.00	−678,650,182.00	369,323,264.00	86,751.08	73.33	3,364,459,482.00	1,551,398,922.00	161,485.55
HU	0.00	0.00	10,147,738.00	91,853.16	50.00	932,612,037.00	409,821,324.00	163,391.22
IE	−6.44	−2,463,000,000.00	0.00	525,764.94	49.74	14,642,000,000.00	5,379,496,576.00	751,081.94
LU	−3.04	−779,702,734.00	0.00	511,342.41	81.96	914,339,800.00	443,556,817.00	512,180.41
PT	−15.57	0.00	1,845,133,926.00	110,196.23	68.18	7,505,375,116.00	3,216,915,642.00	112,844.39
RO	−11.42	0.00	117,966,220.00	58,882.71	37.33	1,667,053,033.00	689,427,538.00	140,152.39
UK	0.00	−1,284,000,000.00	0.00	0.00	59.07	5,727,973,312.00	2,356,033,638.00	256,424.84

* The acronyms for each MS are: BE—Belgium; CY—Cyprus; CZ—the Czech Republic; DE—Germany; DK—Denmark; FR—France; GR—Greece; HU—Hungary; IE—Ireland; LU—Luxemburg; PT—Portugal; RO—Romania; SE—Sweden; SK—Slovakia; UK—the United Kingdom

Source Authors' own elaboration

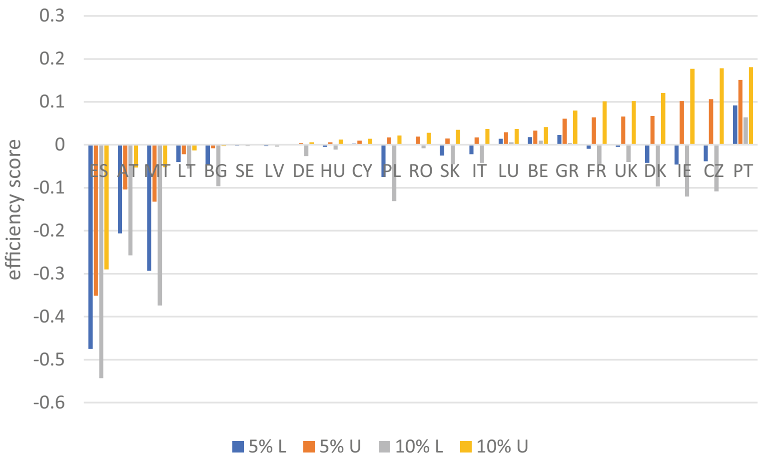


Fig. 7 Lower and upper limits for the value loss, for each MS. *Source* Authors' own elaboration

5 Conclusions and Further Research

The primary goal of this study was to examine the efficiency of ERDF deployment for LCE assistance in 23 EU countries. To tackle this challenge, we propose a two-stage VBDEA approach. The VBDEA model is used in the initial phase to compute each MS's efficiency score. In the second step of the analysis, data were gathered on the required changes to close any gaps between inefficient MS and their benchmarks. Differently from other methodologies used in analogous scenarios, the VBDEA approach is especially significant for MA since it allows evaluating all the MS (either efficient or inefficient) under examination at a single level, assisting in the diagnosis of the causes for their (in)efficiency. Furthermore, because it depends on value functions to convert the DMs' preference information, this approach is straightforward in dealing with null and negative criteria.

The following are the responses to our primary research questions.

RQ1: "What are the main reasons for the (in)efficiency in the utilization of ERDF granted to promote an LCE in EU countries?"

The factors most valued to attain the higher efficiency level possible are "Eligible cost decided" and "Total eligible spending". Moreover, only 4 out of the 10 most efficient MS (Bulgaria, Spain, Italy, and Poland) consider "GHG reduction" to be a key factor for attaining efficiency. The more significant adjustments required to attain efficiency for inefficient MS should be performed in terms of "GHG reduction" and "EU co-financing". This implies that inefficient MS should be concerned about both the selection of initiatives to reduce GHG emissions and their reliance on EU funding.

RQ2: "Which countries were considered benchmarks during the last programmatic timeframe?"

The four most frequently elected benchmarks were Slovakia (8 times), followed ex aequo by Austria and Malta (5 times), and then by Spain (4 times).

RQ3: "Which MS performs better in terms of robustness?"

Spain, Austria, Malta, Lithuania, Bulgaria, and Sweden are robustly efficient (in decreasing order) for both data perturbations used (5 and 10%). Contrastingly, Portugal, Belgium, Luxembourg, Greece, and Cyprus are robustly inefficient (in decreasing order) for both tolerances. Spain is by far the most robust MS in terms of efficiency. Slovakia, which was most often selected as a benchmark is just potentially efficient for all data perturbations.

In conclusion, our results indicate that many EU MS that have effectively deployed renewable sources (see, for example, Germany (inefficient), Spain (efficient), France (inefficient), and Italy (efficient)), when efficient in the application of ERDF devoted to an LCE, benefit from the reduction of GHG emissions to achieve their performance, and when inefficient, do not need to even farther reduce GHG emissions towards becoming efficient, thus being regarded as over users of these types of Funds available. These MS see engagement in LCE (namely, increased renewable deployment) as an economic and political opportunity that allows them to diversify their energy supplies while simultaneously reducing energy imports. Most of these MS are in Western European countries, where they have higher GDP and better-developed energy markets as well as advanced infrastructures. Furthermore, these MS employ a significant portion of their workforce in the renewable power business, which offers them financial advantages amid rising taxes and levies (Pérez et al., 2019). It is also interesting to notice a positive change in two Visegrad MS (Poland and Slovakia), which have generally been opposed to an LCE changeover. Indeed, the Visegrad MS developed a coordinated opposition to both the EU renewables regulations and the EU power market changes (Pérez et al., 2019).

It is also worth noting that these MS are very vulnerable to energy supply disruptions, are particularly dependent on oil, and frequently rely on Russia as a sole source, as well as being located on the EU's periphery. As a result, this change of stance regarding the implementation of an LCE, notably in Poland, Slovakia, and other Eastern EU MS such as Latvia, Lithuania, and Bulgaria, may be partly due to the occupation of Crimea in 2014 (our study covers the period of 2014–2020). Other MS, such as Romania, Hungary, and the Czech Republic, could not manage to effectively apply these funds. Given these findings, the EU must continue to push policies that ensure economic advantages from spending in an LCE, particularly for MS with limited funding.

Acknowledgements This work has been funded by European Regional Development Fund in the framework of Portugal 2020—Programa Operacional Assistência Técnica (POAT 2020), under project POAT-01-6177-FEDER-000044 ADEPT: Avaliação de Políticas de Intervenção Co-financiadas em Empresas. INESC Coimbra and CeBER are supported by the Portuguese Foundation for Science and Technology funds through Projects UID/MULTI/00308/2020 and UIDB/05037/2020, respectively.

References

Ali, A. I., Lerme, C. S., & Seiford, L. M. (1995). Components of efficiency evaluation in data envelopment analysis. *European Journal of Operational Research, 80*(3), 462–473. https://doi.org/10.1016/0377-2217(94)00131-U

Bąk, I., Barwińska-Małajowicz, A., Wolska, G., Walawender, P., & Hydzik, P. (2021). Is the European Union making progress on energy decarbonisation while moving towards sustainable development? *Energies, 14*(13), 3792. https://doi.org/10.3390/en14133792

Chen, Y., Liu, S., Wu, H., Zhang, X., & Zhou, Q. (2020). How can Belt and Road countries contribute to glocal low-carbon development? *Journal of Cleaner Production, 256*, 120717. https://doi.org/10.1016/j.jclepro.2020.120717

European Commission. Guidance Document on Monitoring and Evaluation. European Cohesion Fund, European Regional Development Fund. Concepts and Recommendations. (2014). Available online: https://ec.europa.eu/regional_policy/sources/docoffic/2014/working/wd_2014_en.pdf (Accessed on 19 December 2021)

Gouveia, M. C., Dias, L. C., & Antunes, C. H. (2008). Additive DEA based on MCDA with imprecise information. *Journal of the Operational Research Society, 59*(1), 54–63. https://doi.org/10.1057/palgrave.jors.2602317

Gouveia, M. C., Dias, L. C., & Antunes, C. H. (2013). Super-efficiency and stability intervals in additive DEA. *Journal of the Operational Research Society, 64*(1), 86–96. https://doi.org/10.1057/jors.2012.19

Gouveia, M. C., Henriques, C. O., & Costa, P. (2021). Evaluating the efficiency of structural funds: An application in the competitiveness of SMEs across different EU beneficiary regions. *Omega, 101*, 102265. https://doi.org/10.1016/j.omega.2020.102265

Henriques, C., Viseu, C., Trigo, A., Gouveia, M., & Amaro, A. (2022a). How efficient is the cohesion policy in supporting small and mid-sized enterprises in the transition to a low-carbon economy? *Sustainability, 14*(9), 5317. https://doi.org/10.3390/su14095317

Henriques, C., Viseu, C., Neves, M., Amaro, A., Gouveia, M., & Trigo, A. (2022b). How efficiently does the EU support research and innovation in SMEs? *Journal of Open Innovation: Technology, Market, and Complexity, 8*(2), 92. https://doi.org/10.3390/joitmc8020092

Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value trade-offs*. Cambridge University Press.

Levy, C. (2010). A 2020 low carbon economy a knowledge economy programme report. The Work Foundation. Retrieved January 02, 2022, from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.620.4765&rep=rep1&type=pdf

Liu, X., & Liu, J. (2016). Measurement of low carbon economy efficiency with a three-stage data envelopment analysis: A comparison of the largest twenty CO_2 emitting countries. *International Journal of Environmental Research and Public Health, 13*(11), 1116. https://doi.org/10.3390/ijerph13111116

Meng, M., Fu, Y., & Wang, L. (2018). Low-carbon economy efficiency analysis of China's provinces based on a range-adjusted measure and data envelopment analysis model. *Journal of Cleaner Production, 199*, 643–650. https://doi.org/10.1016/j.jclepro.2018.07.219

Pérez, M. D. L. E. M., Scholten, D., & Stegen, K. S. (2019). The multi-speed energy transition in Europe: Opportunities and challenges for EU energy security. *Energy Strategy Reviews, 26*, 100415. https://doi.org/10.1016/j.esr.2019.100415

Zha, J., He, L., Liu, Y., & Shao, Y. (2019). Evaluation on development efficiency of low-carbon tourism economy: A case study of Hubei Province, China. *Socio-Economic Planning Sciences, 66*, 47–57. https://doi.org/10.1016/j.seps.2018.07.003

Zhang, J., Zeng, W., Wang, J., Yang, F., & Jiang, H. (2017). Regional low-carbon economy efficiency in China: Analysis based on the Super-SBM model with CO_2 emissions. *Journal of Cleaner Production, 163*, 202–211. https://doi.org/10.1016/j.jclepro.2015.06.111

Zhang, Y., Shen, L., Shuai, C., Tan, Y., Ren, Y., & Wu, Y. (2019). Is the low-carbon economy efficient in terms of sustainable development? A Global Perspective. *Sustainable Development, 27*(1), 130–152. https://doi.org/10.1002/sd.1884

Zhang, Y., Wang, W., Liang, L., Wang, D., Cui, X., & Wei, W. (2020). Spatial-temporal pattern evolution and driving factors of China's energy efficiency under low-carbon economy. *Science of the Total Environment, 739*, 140197. https://doi.org/10.1016/j.scitotenv.2020.140197

Ex-post Assessment: Meta-analysis

Firm Competitiveness, Specialisation, and Employment Growth: Territorial Level Relationships

Federico Fantechi and Ugo Fratesi

Abstract The concept of competitiveness is today a central element for regional development, European cohesion policies and smart specialisation strategies. Despite being born for firm-level analyses, competitiveness is indeed commonly used at the territorial level, mainly at the regional or urban scale, normally measured with different composite structural indicators. However, since territorial competitiveness is unevenly distributed in space, territorial units smaller than a full NUTS-2 region might be differently competitive and hence suited to implement differentiated cohesion policies and smart specialisation strategies. To test the hypothesis that these firm-level indicators can characterize the intraregional differences in aggregate performance, the paper sets up a meta-analysis framework between these indicators and structural indicators (employment growth and specialisation index) measured at the NUTS-3 level. For the meta-analysis at this novel intraregional level, the paper exploits the Lombardy region as a case study. Lombardy is well suited for the aims of this paper, being a large and competitive European region, whose territory—as well as its labor market—is highly differentiated, from peripheral and mountainous areas to many medium and small cities, second-tier large cities and a large metropolitan area—the city of Milan. All these territories are characterized by different economic and social vocations, but all share the same regional administration. The results of the meta-analysis show that firm-level indicators correlate with the aggregate performance of regions and that the structural measures selected can characterize different territories in different conditions. Hence, the competitiveness of firms seems to translate into aggregate territorial performance at small spatial scales. This implies that territorial specificities are also relevant inside regions and should be considered in designing regional policy interventions, such as those of the Smart Specialisation Strategy (S3).

F. Fantechi (✉) · U. Fratesi
Politecnico di Milano Department of Architecture, Built Environment and Construction Engineering, Piazza Leonardo da Vinci, 32 20133 Milano, Italy
e-mail: federico.fantechi@polimi.it

U. Fratesi
e-mail: ugo.fratesi@polimi.it

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_6

Keywords Smart specialisation · Firm productivity · Employment growth

1 Introduction

With the programming period 2014–2020, European Union (EU) Cohesion Policies (CP) introduced the key concept of smart specialisation (Foray, 2015), which further focused EU cohesion policies around the two main elements of innovation and territorial competitiveness, fitting smart specialisation as an ex-ante condition for receiving support from European structural and investment funds (Landabaso, 2014; Mccann & Ortega-Argilés, 2013). A key aspect of these smart specialisation strategies is the centrality of the context in which they are implemented. Indeed, following the growing emphasis gained by place-based policies (Barca et al., 2012), the proposed reforms aim to better link institutions, policies and incentives around and with the territorial context and evolutionary trends of the regions.

Connectedly with the rise of smart specialisation strategies, another concept returned to centrality in the allocation and design of regional policies: territorial competitiveness. Despite being originally conceived as firm-related, the concept of competitiveness has also been applied to analyze territories since the early 1990s (Porter, 1990); due to its direct link with the capacity for production—either at the firm or territorial level—today, the concept of competitiveness is a common element for policy design, especially regarding policy programs aimed at reducing the productivity gaps.

Policies built with these elements at their core are designed and allocated with the intent of nurturing and supporting regions and territories that are best competing in the international market but also to help lagging or underdeveloped regions and territories "in order to build competitive advantage by developing and matching research and innovation" (REGULATION (EU) No 1303/2013 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 December 2013, article 2).

Smart specialisation strategies are inherently territorialized, and their main strength is that they aim to improve economic performance and development paths by fostering and exploiting local knowledge and territorial capital. However, when pragmatically observing how these territorial aims can be achieved, a "mismatch" emerges between the allocation and effective implementation of these policies.

Indeed, regional and cohesion policies are commonly allocated at the NUTS-2 level[1]; likewise, most relevant measures and indicators are also aggregated at that level. Most notably, this is the case for the Regional Competitiveness Index (RCI) computed by the EU Commission at the NUTS-2 level accounting for multiple characteristics of a region and its industrial structure into a single measure comparable between regions.

[1] NUTS stands for Nomenclature of Territorial Units for Statistics, it is the hierarchical classification of European territorial units.

On the other hand, these policies are to be implemented at the territorial level unevenly over the regional territory. As abundantly shown by prominent scholars in regional science, a large set of influencing factors are highly territorialized and unevenly distributed in space (e.g., infrastructure, human capital, skilled workers and quality institutions) (Dierickx & Cool, 1989; Maskell & Malmberg, 1999). Moreover, firms and economic activities (both related and nonrelated) are not evenly distributed over the regional territory.

All these elements, commonly called territorial capital (Camagni, 2009; Fratesi & Perucca, 2019), are highly distributed in the regional territory and influence it at a much smaller scale than NUTS-2.

This paper argues that the set of information and instruments available for the allocation and the design of these policies may not match the territorial level on which they take place. Therefore, a different instrument—measuring territorial competitiveness at the subregional level—is needed to better inform the design and implementation of smart specialisation strategies.

By means of a novel methodology and firm-based territorial analyses correlating specialisation, employment growth and territorial competitiveness, this paper aims to show how it is possible to produce subregional measures of territorial competitiveness—exploiting firm-level data—providing territorial information with varying territorial units on which to design these place-based policies.

Focusing on a single NUTS-2 region, the Lombardy region in Italy, the paper first presents how intraregional territorial competitiveness can be measured via firm-level data. Then, a dynamic meta-analysis shows how these measures correlate with specialisation and territorial growth and development.

In its concluding remarks, the paper argues that—as shown by the results of the meta-analysis—this intraregional measure of territorial competitiveness can provide vital territorial information for the design of place-based policies and, in the context of Smart specialisation Strategies (S3) and Regional Cohesion Policies can be a very useful instrument to implement—rather than replace—aggregate indicators such the Regional Competitiveness Index.

2 Smart Specialisation Strategies and Territorial Competitiveness: A Missing Link Between Theory and Practice

2.1 S3 and Territory

In the contemporary EU policy debate, smart specialisation strategies are the central node of many policy programs and designs. Indeed, the agenda resulting from this paradigm-shifting concept was in the programming period 2014–2020 an ex-ante condition for receiving support from EU structural and investment funds in TO1

(Landabaso, 2014) and will remain fundamental in the programming period 2021–2027.

At their core, smart specialisation strategies assume that the context in which firms operate not only matters but can also be the main driver of the technological evolution of innovation systems. Existing strengths but also untapped potentials of territories can be exploited to foster—through these agendas—growth and maximize the development opportunities of territories and regions. Despite the name, the smart specialisation agenda is not intended to encourage sectoral specialisation but rather to foster diversification around a core set of activities and generate new specialties and opportunities for local concentration and agglomeration of resources and competences in these domains' (Foray, 2015 p.1).

From this perspective, the context—with its local knowledge networks, trade links, spillovers and everything else that today is considered a key element of the related variety (Boschma & Iammarino, 2009; Boschma et al., 2012; Frenken et al., 2007; Neffke et al., 2011)—is considered the existing structure on which to develop a 'diversified' specialisation (Grillitsch et al., 2018) and to foster related explorative, research activities (Foray, 2014).

It is important to note that although smart specialisation strategies do not target specific territories and regions, most positive examples of such strategies are located in structurally and economically strong regions (Foray, 2015). It is clear that by heavily relying on locally existing strengths and opportunities, the effectiveness of these strategies is largely impacted by the development path and industrial past of the region. This is where a missing link emerges; while both policy frameworks and policy actors have switched already their perspectives, from a regional to a – smaller – territorial one, there still is a lack of tools and support instruments (such as the European Regional Competitiveness Index, which only considers the regional level) available to a smaller level than the NUTS-2.

2.2 Case Study Description

For every empirical study, defining and selecting a well-suited case study is a key step.

In this case, it is important to select NUTS-3 areas belonging to the same NUTS-2 region because only in this way the institutional framework will be the same for all. At least this is what happens in Italy, the country from which data come from, where NUTS-2 regions are endowed with large autonomy.

We select the most competitive region of the country, Lombardy, which is also the largest in terms of population, territory, and total GDP.[2] The region has just one smart specialisation strategy although it is composed of many different territories

[2] Lombardy not only consistently scores higher than the rest of the country on the RCI (Regional Competitiveness Index) but also holds a higher GDP per capita of €39,200 in 2018, compared to the average of €29,700 in Italy and €31,000 in the EU (Eurostat, 2020).

with their geographical and economic specificities. In particular, this region includes one large metropolitan area the city of Milan as well as medium-sized cities and more peripheral areas both in the plains and in the mountains. Overall, Lombardy holds a large territory and is the most populated region in the country, almost doubling the population of the second largest region with almost 10 Million inhabitants (Istat, 2022).

3 Data and Methodology

3.1 Measuring Territorial Competitiveness Using Firm-Level Data

The competitiveness of territories and that of the companies located within those territories are intrinsically connected. Indeed, the competitive capacity of a particular firm is influenced by three sets of factors: (i) the characteristics of the individual firms; (ii) the dynamics of the industrial sectors; and (iii) a large set of territorial elements and characteristics which, taken together, are called territorial capital.

Exploiting this general assumption, a "two-step" matching design (Rosenbaum & Rubin, 1985) is implemented to isolate the differential effects on the competitiveness of firms produced overall by those elements known as territorial capital (Camagni, 2009; Fratesi & Perucca, 2019). If in fact, two firms in the same industrial sector share similar characteristics—being different only in terms of their location—resulting differences in terms of competitiveness between those firms will be due to the external conditions in which they operate. By aggregating firms based on their location in one of the 12 NUTS-3 provinces inside the Lombardy region, a "two-step" matching design is implemented to separately control for industrial dynamics and individual firm characteristics. The produced differentials can easily be employed to proxy internal differences in territorial competitiveness.

This counterfactual workflow, recently proposed by (Fantechi & Fratesi, 2022), has a number of advantages, especially over the use of composite indices. Indeed, it employs firm-level microdata instead of administrative statistical or census data; the availability of firm-level microdata is constantly growing and, especially for European countries, today several databases are available detailing firms' master and balance information for almost the last two decades. Moreover, the workflow is quite flexible, allowing easy variation in both the level of analysis and the area of study and allowing for both static and dynamic enquiries. Nevertheless, the current formulation of the workflow, despite allowing for a certain degree of freedom and being able to control and isolate from industrial sectorial dynamics and differences in firms' characteristics, is not able to differentiate between the first and second "nature" of territorial capital.

The counterfactual strategy implemented to measure intraregional territorial differences in territorial capital is a "two-step" matching design (Rosenbaum & Rubin, 1985).

Each of the two steps of the strategy is designed to control different influencing factors of firms' competitiveness and thus isolate the territorial effect. The first step in the matching design consists of an exact match for the industrial sector in which the firms operate. Indeed, the industry in which a specific firm operates is probably the most influential single aspect to account for. Firms operating in different markets may not only have very different production margins, market sizes and organizational requirements but may also differ in terms of growth and dynamic opportunities. The overall effect of being part of different industrial sectors is considered by matching firms sector by sector using the NACE 4-digit classification and the 22 categories following the STAN industry list ISIC rev.4 classification (Horvát & Webb, 2020).

Based on this fine classification of the main industrial sector in which a firm operates, firms are matched, and their performance compared, only with other firms in the same class.

The second step of this matching design is composed of propensity score matching (via a probit function with a caliper of 0.05) to control for past trends and specific firm characteristics. In this second step of the strategy, the aim is to isolate the differentials in firms' performance—based on the NUTS-3 territory in which they are located—from the influence of specific firms' characteristics. To do so, the probit function controls our data for several characteristics of firms to only compare firms in each industrial sector only with similar firms (in the same industrial sector) located elsewhere.

Several are the characteristics selected for this operation:

The ***age of the firm*** is accounted for via a discrete variable recording the number of years passed from the registered incorporation of the firm.

Being a ***beneficiary of public policy*** interventions or not is indicated via a dummy that identifies those firms that received some kind of public assistance in the years prior to the research.

Whether firms have a ***cooperative status*** is accounted for with a dummy variable identifying those firms incorporated as cooperatives.

Involvement in international markets is again accounted for with a dummy variable (due to availability of data) identifying those firms who self-report ***export activities***.

Firm size is indicated by a discrete variable recording the number of employees.

The ***reliance on immaterial assets*** by firms is accounted for by means of a continuous variable measuring the share of immaterial assets (over total assets) declared by a specific firm.

Finally, ***financial position*** is accounted for by a continuous variable measuring the ratio of debts to total gross earnings.

All these variables are computed on firm-level, self-reported, yearly data recovered from the AIDA database (Bureau van Dijk, n.d.).

In this way, it is possible to compare firms present in one of the 12 provinces of Lombardy with other firms which belong to the same sector and are structurally similar but are located in a different province. The differentials arising will depend on territorial capital of the provinces.

Three indicators are used as measures of productivity and profitability.

Labor productivity: computed as the ratio between Value Added and number of Employees (Aguiar & Gagnepain, 2017; Bhattacharya & Rath, 2020; Falciola et al., 2020; Laureti & Viviani, 2011; Nemethova et al., 2019).

Total Factor Productivity: computed as the residual of a Solow production function (Solow, 1956) based on Value Added and calculating the capital stock at the firm level using the Perpetual Inventory Method (PIM) (Gal, 2013), thus also including the firm's capitals and capitalization in the computation (Albanese et al., 2020; Ciani, Locatelli, & Pagnini, 2018; Gal, 2013; Lasagni et al., 2015).

Profitability: measured as a ratio of EBITDA on Turnover, also known as ROA, Return On Assets (Aguiar & Gagnepain, 2017; Akimova, 2000; Bharadwaj, 2000; Bramanti & Ricci, 2020).

As a final control, specific to the analyzed case study, the research also accounts for an eventual "sorting effect" in the localization selection by firms. Indeed, large cities, especially large metropolitan areas, are exceptionally more attractive to firms than other territories, producing results—in terms of firms' productivity—often on a different scale. This is due not only to higher stocks of territorial capital but also to being a "place on the map" (i.e., branding opportunities, name recognition) (Wheeler, 2001) and providing unique opportunities. To avoid the possible confusion generated by this sorting effect, a simple restriction is implemented in the matching design to account for this effect without affecting or penalizing firms located in different territories: the province of Milan (which is mostly composed of the metropolitan area of Milan, the only truly "big" city in the region) is compared with the rest of the region to calculate the competitiveness differential for firms of being located there; conversely, when matching firms from the other provinces, firms located in the province of Milan are excluded from the computation.

The time span of the analysis includes a period of ten years, between 2009 and 2019. Two main types of data are required for the analysis: i) firm-level balance sheet data, provided by AIDA (Bureau van Dijk, n.d.); territorial, administrative-level data provided by Istat (15th Italian Census: ISTAT, 2011) and ASIA (National registry of Firms: ISTAT, 2020).

As shown in (Fantechi & Fratesi, 2022), the produced differential can easily and effectively be employed to identify and characterize internal differences in territorial capital. However, this is not the focus and aim of the present paper; the results of the described two-step matching design will indeed serve as input data for a meta-analysis connecting them to structural indicators of specialisation and regional growth and development.

3.2 Meta-Analysis

The final objective of the analysis is to show how the proposed measure of territorial competitiveness—compiled at the intraregional level from firm-level data—correlates to established measures of regional growth and specialisation. Territorial competitiveness plays a key role in today's policy design and implementation and has often been positively correlated with a positive impact on regional growth and development. While this correlation has been shown to exist at the regional level, where territorial competitiveness is measured via composite indicators, this paper wants to show that the same correlation also stays true at the intraregional level. Moreover, by measuring territorial competitiveness as a residual of the firms' competitiveness differential based on their location, the paper also provides a novel—and quite adaptable—methodology to measure territorial competitiveness; showing how this measure of territorial competitiveness correlates with—connected—more established measures of regional growth and specialisation will provide additional data in favor of the use of this indicator and more detailed territorial information on which design more effective cohesion and industrial policies.

While the analysis per se—consisting of a set of multiple regressions—is quite straightforward, it is worth describing in more detail both the selection of indicators and the data operations prior to inputting them into the models.

3.3 Specialisation Indicators and Data Preparation

To correlate the competitiveness of territories inside a specific NUTS-2 to their territorial growth and development, a viable measure of growth at the NUTS-3 level must be identified. Territorial growth, which is commonly measured at the NUTS-2 level, is a complex concept encompassing various dimensions from individual well-being, social inclusion, economic prosperity, and structure to access to services and institutions. Many of the metrics employed are directly recorded or measured from national and supra-national statistical offices, mostly at the NUTS-2 level. Considering the scarcity of such measurements at the NUTS-3 level and that the aims of this paper are directly connected to the specific dimension of territorial economic prosperity, only one measure of territorial growth has been selected: employment growth. Differently from GVA, this variable is able to account for the territorial effects of economic aspects in a way which also considers its social consequences, in terms of employment (Fratesi & Rodriguez-pose, 2016).

For this analysis, the growth in employment is measured as the relative change between 2007 and 2019; this is possible thanks to data from ASIA (The Italian registry of active firms) reporting the total number of employed workers in each industrial sector at the municipal level.

The relative change in employment is calculated for each Nace 2-Digit division (aggregating less relevant and numerous sectors) at the NUTS-3 provincial level for the whole region. The same unit of analysis is also employed for the computation of differential.

Finally, output data—from the analysis of territorial competitiveness performed employing firm-level data—need to be processed before imputing them into the model.

As described in the first part of the methodological paragraph, territorial competitiveness is measured via firm-level data producing territorial competitiveness differentials for each of the analyzed territories.

To improve the reliability and explanatory power of the meta-analysis, output data are processed and discretized before imputing them. Indeed, it is important to consider that the produced counterfactual results—which, after being processed, will become input data for the meta-analysis—are normalized territorial differentials coefficients of competitiveness measured via firm-related data. The paper is interested in the territorial-level relationship between the differences in competitiveness detected by the counterfactual strategy and different territorial trends in terms of employment growth. Directly imputing the coefficients in the models would not provide additional information on this relationship, while at the same time, it would produce a much more complicated and less reliable model. For this reason, before imputing, each coefficient has been discretized taking one of three possible values: (1) "Not significant" for those coefficients which are, regardless of the sign, not statistically significant; (2) "Positive" for those coefficients which are both positive in sign and statistically significant; (3) "Negative" for those coefficients which are both negative in sign and statistically significant.

Additional controls are included in the analysis to provide more robustness to the results. Both industrial sector controls and spatial controls (NUTS-3 level) are included; moreover, "specialized", and "nonspecialized" territories are identified before computing the analysis: for each observation, a specific industrial sector in a specific province (NUTS-3), we identified whether it is "more specialized than average" or "less specialized than average" by exploiting sectoral employment location quotients.

4 Results and Discussion

With the aim of bridging the gap between existing structural indicators of territorial (regional) competitiveness and the need for more territorialized measures of competitiveness to inform the design and implementation of policies, the paper presented—following (Fantechi & Fratesi, 2022)—a counterfactual workflow to measure differentials of territorial competitiveness at the subregional level.

This is done by employing firm-level data with a 2-step matching strategy: the first step eliminates the heterogeneity produced by firms being part of different industrial sectors with an exact match, ensuring that firms operating in a specific sector are only matched and compared with firms in the same exact sector. With propensity score matching over individual firms' characteristics, the second step controls for the different conditions in which the firm operates (e.g., size, initial production capacity, different assets reliance, financial position) so that firms are only compared to similar firms located elsewhere. As argued in (Fantechi & Fratesi, 2022), the produced differentials are indicators of differences in territorial competitiveness produced by the different distribution and availability of territorial capital inside the region.

The main aim of this paper is, then, to show and test the correlation between the produced differentials and established measures of territorial competitiveness measured at the same territorial level.

Before inputting data for the meta-analysis, firm-level territorial differentials produced with the counterfactual strategy are processed and discretized as discussed in the previous section. ATTs from all three indicators (labor productivity, total factor productivity and profitability) are calculated and inputted, measuring three different—and connected—sides of the competitiveness of firms, both in static form and dynamic one (for a total of six indicators).

The meta-analysis is performed by means of multiple linear regressions on the change in employment. Figure 1 reports the results of such meta-analysis where each indicator takes value "1" if the specific computed differential is statistically significant and positive and 0 otherwise (significant and negative, or not significant). This is done to both provide a more readable output and simplify the analysis to better show the correlation between change in employment and the computed differentials.

Figure 1 shows a number of interesting results concerning the regional specialisation and especially the openness indicators. The table is organized in columns with different regression models where alternative specifications are presented.

CANGHE IN EMPLOYMENT	(1) Specialization Only	(2) Static Only	(3) Dynamic Only	(4) Both	(5) NUTS-3 Controls	(6) No Milan	(7) Specialized Only	(8) Specialized and Average	(9) Not Specialized and Average	(10) Not Specialized Only
Specialization										
More than Average	-0.0442*	-0.0433*	-0.0508**	-0.0504**	-0.0422*	-0.0548**		-0.0512**		
	(0.0239)	(0.0229)	(0.0242)	(0.0236)	(0.0247)	(0.0242)		(0.0236)		
Less than Average	-0.0103	-0.00858	-0.0157	-0.0162	-0.00541	-0.0182			-0.0228	
	(0.0241)	(0.0234)	(0.0257)	(0.0253)	(0.0263)	(0.0267)			(0.0260)	
Static indicators										
Positive ATT Labor Prod		0.0721^		0.0650^	0.0642^	0.0833^	0.0505	0.0809^	0.0607	-0.0341
		(0.0455)		(0.0445)	(0.0454)	(0.0615)	(0.0558)	(0.0540)	(0.0575)	(0.0601)
Positive ATT TFP		-0.00345		0.00927	0.0177	0.0252	0.00719	0.00280	0.00664	0.0218
		(0.0201)		(0.0187)	(0.0242)	(0.0238)	(0.0549)	(0.0205)	(0.0211)	(0.0395)
Positive ATT ROA		0.00394		0.00410	0.00597	0.00402	0.0678	-0.00530	-0.00547	0.0175
		(0.0187)		(0.0189)	(0.0205)	(0.0195)	(0.0769)	(0.0252)	(0.0207)	(0.0344)
Dynamic indicators										
Positive ATT Labor Prod			0.00416	-0.00213	-0.000631	-0.00783	0.0209	0.0227	-0.0140	-0.148**
			(0.0264)	(0.0280)	(0.0269)	(0.0286)	(0.109)	(0.0278)	(0.0298)	(0.0666)
Positive ATT TFP			0.0580***	0.0481**	0.0520**	0.0554*	0.0972*	0.0654***	0.0185	0.0171
			(0.0223)	(0.0230)	(0.0250)	(0.0304)	(0.0554)	(0.0214)	(0.0302)	(0.0953)
Positive ATT ROA			0.0995*	0.101*	0.0971*	0.103*	0.0928	0.121^	0.103^	0.101*
			(0.0594)	(0.0572)	(0.0543)	(0.0571)	(0.101)	(0.0799)	(0.0710)	(0.0575)
Nace Sector Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	0.0366	0.0295	0.0321	0.0265	-0.0105	0.0256	-0.0593	0.0303	0.0545	-0.0165
	(0.0579)	(0.0590)	(0.0567)	(0.0582)	(0.0701)	(0.0643)	(0.0447)	(0.0918)	(0.0818)	(0.0239)
Observations	260	260	260	260	260	238	54	183	206	77
R-squared	0.476	0.492	0.506	0.519	0.547	0.527	0.684	0.574	0.520	0.583

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, ^ p<0.15

Fig. 1 Meta-analysis results *Source* Authors' own elaboration

The left side of the table presents relations with the full sample and first shows that territories with higher than the average specialisation in their industries have produced a worse performance in terms of employment.

On the contrary, these territories where firms are more competitive than their counterparts, usually have a significantly better performance. This is true in particular for the indicator of total factor productivity; this is the most significant coefficient and shows that in those places where firms are more innovative, additional jobs are created.

As expected, there is also a positive coefficient for labor productivity showing that where labor is more productive, the firms of these territories react by hiring additional workers.

There is however a third result which is somehow counterintuitive. When firms increase their profitability then they also create jobs. This is demonstrated by the positive and significant coefficient of ROA. At least in these advanced regions, the profitability of firms is not in contrast with the creation of jobs, and the most competitive firms are usually innovative so that they produce returns for their investors and at the same time additional work for their local communities.

The right part of the table presents results for regressions on different sub-samples to see if the effects detected depend on specialisation.

It shows that the positive impact of total factor productivity is mostly present in areas of specialisation. The positive effect of profitability is instead present in all areas but the most specialised.

5 Conclusions and Further Research

In this paper, we identified a 'missing link' between the level at which smart specialisation programs are assigned and designed and the level at which they are applied. The key element in building this research is indeed the mismatch between the availability of tools and instruments for policy programming and design and the territories in which the policies are to be implemented.

Focusing on territorial competitiveness—central element and aim of Smart specialisation Strategies—this missing link is evident: tools, indicators, and indices of territorial competitiveness (especially those produced by EU's institutions) provide information at the regional (NUTS-2) level; in order to maximize both the efficiency of policy design and the effectiveness of policy implementation, Smart Specialisation Strategies can greatly benefit from more territorialized measures and indicators of territorial competitiveness. In a recent publication, Fantechi and Fratesi (2022) developed an adaptive framework to measure differentials of territorial competitiveness inside a region. This framework, presented in Sect. 3 of this paper, employs firm-level data to provide territorialized firms' competitiveness indicators by isolating and controlling the effects of industrial sectorial dynamics and firms' individual characteristics. According to the authors, the main feat of the framework is its adaptability to different levels of analysis, being smaller administrative units (as performed for the analysis in this paper) or specific geographical areas. Information provided with this framework is not intended to replace existing indicators and indices of territorial competitiveness; rather, to implement them with information they are not able to provide to help the design and implementation of smart specialisation strategies.

The main result of this article was to show the relation between these territorialized firms' competitiveness measures, specialisation, and territorial growth (growth in employment) to test and validate the indicators as a valuable tool to measure subregional differentials of territorial competitiveness and performance.

Following the outlined framework, multiple territorialized firms' competitiveness indicators of both productivity and profitability of firms are computed at the provincial (NUT-3) administrative level. The results provided in this paper show the interesting potential of this tool. One main result is that positive territorialized differentials of productivity (TFP) and profitability (ROA) correlate positively with larger local growth in employment over the same period. Interestingly, the third selected indicator of territorialized firms' competitiveness, labor productivity, shows a lower correlation.

Taken together, these results show that the proposed framework can help individuate increasingly competitive territories inside the region and can also provide an indication of which elements of firms' competitiveness local territories are able to provide better support to local employment growth.

The territorial-level relation between firms' competitiveness—through their territorialized indicators—and the growth in employment shows the possible relevance of the framework developed by Fantechi and Fratesi (2022), not only as a research tool but also as a tool for policy design providing relevant information on the competitiveness of territories inside a region.

The limits and shortcomings of this approach are multiple and represent the main reason why the proposed framework is not intended to replace existing measures and indicators but, rather, to integrate them. Some of these limits are inevitable due to the framework itself; the produced territorialized firms' competitiveness indicators are territorial differentially produced with a counterfactual 2-step matching, meaning that they are relative measures rather than absolute. They correctly represent internal territorial differences, but to provide a correct interpretation, they are related to the overall context and dynamics of the region; results emerging from this analysis cannot be generalized and directly applied to other European regions without considering the relevance and specificities of different regional dynamics and characteristics. Finally, other smaller limits are due to the availability of data and information, specifically firm-level data; firm-level balance sheet data available today are, despite being a great resource, still partially lacking in precision and completeness. This clearly represents a limitation to this study (as well as other studies employing the same data) but a limitation that is destined to fade in the coming years as the database becomes more complete and the capacity for data gathering and production refines.

References

Aguiar, L., & Gagnepain, P. (2017). European cooperative R&D and firm performance: Evidence based on funding differences in key actions. *International Journal of Industrial Organization, 53*, 1–31. https://doi.org/10.1016/j.ijindorg.2016.12.007

Akimova, I. (2000). Development of market orientation and competitiveness of Ukrainian firms. *European Journal of Marketing, 34*(9/10), 1128–1148. https://doi.org/10.1108/03090560010342511

Albanese, G., de Blasio, G., & Locatelli, A. (2020). Does EU regional policy promote local TFP growth? Evidence from the Italian Mezzogiorno. Papers in Regional Science, 0–3,. https://doi.org/10.1111/pirs.12574

Barca, F., Mccann, P., & Rodríguez-Pose, A. (2012). The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science, 52*(1), 134–152. https://doi.org/10.1111/j.1467-9787.2011.00756.x

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly: Management Information Systems, 24*(1), 169–193. https://doi.org/10.2307/3250983

Bhattacharya, P., & Rath, B. N. (2020). Innovation and firm-level labour productivity: A comparison of Chinese and Indian manufacturing based on enterprise surveys. *Science, Technology and Society, 25*(3), 465–481. https://doi.org/10.1177/0971721820912902

Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography, 85*(3), 289–311. https://doi.org/10.1111/j.1944-8287.2009.01034.x

Boschma, R., Minondo, A., & Navarro, M. (2012). Related variety and regional growth in Spain*. *Papers in Regional Science, 91*(2), 241–256. https://doi.org/10.1111/J.1435-5957.2011.00387.X

Bramanti, A., & Ricci, S. (2020). Structure and performance of the Italian alpine "core": A counterfactual analysis. *Worldwide Hospitality and Tourism Themes*. https://doi.org/10.1108/WHATT-05-2020-0026

Camagni, R. (2009). Territorial capital and regional development. In R. Capello & P. Nijkamp (Eds.), *Handbook of Regional Growth and Development Theories* (pp. 118–132). Edward Elgar.

Ciani, E., Locatelli, A., & Pagnini, M. (2018). Evoluzione territoriale della TFP: analisi dei dati delle società di capitali manifatturiere tra il 1995 e il 2015 (Questioni di Economia e Finanza).

Dierickx, I., & Cool, K. (1989). Asset stock accumulation and sustainability of competitive advantage. *Management Science, 35*(12), 1504–1511. https://doi.org/10.1287/mnsc.35.12.1504

Eurostat. (2020). Regional Innovation Monitor.

Falciola, J., Jansen, M., & Rollo, V. (2020). Defining firm competitiveness: A multidimensional framework. *World Development, 129*, 104857. https://doi.org/10.1016/j.worlddev.2019.104857

Fantechi, F., & Fratesi, U. (2022). Measuring competitiveness differentials inside the same region: A propensity-score matching approach. *Social Indicators Research*, 1–27. https://doi.org/10.1007/s11205-022-02908-7

Foray, D. (2015). *Smart specialisation: Opportunities and challenges for regional innovation policy*. Routledge.

Foray, D. (2014). Smart specialisation. *Smart Specialization* (December). https://doi.org/10.4324/9781315773063

Fratesi, U., & Rodriguez-pose, A. (2016). The crisis and regional employment in Europe. Working Paper, pp. 33–57.

Fratesi, U., & Perucca, G. (2019). EU regional development policy and territorial capital: A systemic approach. *Papers in Regional Science, 98*(1), 265–281. https://doi.org/10.1111/pirs.12360

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies, 41*(5), 685–697. https://doi.org/10.1080/00343400601120296

Gal, P. N. (2013). Measuring Total Factor Productivity at the Firm Level using OECD-ORBIS Recent. OECD Economics Department Working Papers (1049), 1–59.

Grillitsch, M., Asheim, B., & Trippl, M. (2018). Unrelated knowledge combinations: The unexplored potential for regional industrial path development. *Cambridge Journal of Regions, Economy and Society, 11*(2), 257–274. https://doi.org/10.1093/CJRES/RSY012

Horvát, P., & Webb, C. (2020). The OECD STAN Database for industrial analysis: Sources and methods. OECD Science, Technology and Industry Working Papers, OECD Publishing, Paris, 2020/10. https://doi.org/10.1787/ece98fd3-en

ISTAT. (2011). 15° Censimento della popolazione e delle abitazioni 2011. Retrieved from www.istat.it

ISTAT. (2020). Registro statistico delle imprese attive (ASIA). Retrieved from www.istat.it

ISTAT. (2022). Censimento permanente. Retrived from www.istat.it

Landabaso, M. (2014). Guest editorial on research and innovation strategies for smart specialisation in Europe. *European Journal of Innovation Management, 17*(4), 378–389. https://doi.org/10.1108/EJIM-08-2014-0093

Lasagni, A., Nifo, A., & Vecchione, G. (2015). Firm productivity and institutional quality: Evidence from italian industry. *Journal of Regional Science, 55*(5), 774–800. https://doi.org/10.1111/jors.12203

Laureti, T., & Viviani, A. (2011). Competitiveness and productivity: A case study of Italian firms. *Applied Economics, 43*(20), 2615–2625. https://doi.org/10.1080/00036840903357439

Maskell, P., & Malmberg, A. (1999). The competitiveness of firms and regions. "Ubiquitification" and the importance of localized learning. *European Urban and Regional Studies*, *6*(1), 9–25. https://doi.org/10.1177/096977649900600102

Mccann, P., & Ortega-Argilés, R. (2013). Regional studies smart specialisation, regional growth and applications to european union cohesion. *Policy*. https://doi.org/10.1080/00343404.2013.799769

Neffke, F., Henning, M., & Boschma, R. (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography, 87*(3), 237–265. https://doi.org/10.1111/j.1944-8287.2011.01121.x

Nemethova, V., Siranova, M., & Sipikal, M. (2019). Public support for firms in lagging regions-evaluation of innovation subsidy in Slovakia. *Science and Public Policy, 46*(2), 173–183. https://doi.org/10.1093/scipol/scy046

Porter, M. E. (1990). The competitive advantage of nations. *Harvard Business Review*, *68*, 73–93.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39*(1), 33–38. https://doi.org/10.1080/00031305.1985.10479383

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, *70*(1), 65–94. Retrieved from http://qje.oxfordjournals.org/content/70/1/65.short

van Dijk, B. (n.d.). Balance sheet data for Italian Firms. AIDA.

Wheeler, C. H. (2001). Search, sorting, and urban agglomeration. *Journal of Labor Economics, 19*(4), 879–899. https://doi.org/10.1086/322823

The Youth Guarantee in Eastern Europe. A Systematic Review

Gabriela Neagu

Abstract The Youth Guarantee (YG) is one of the most innovative European programs for young people. Given the complexity of the program for the analysis of its effectiveness and efficiency, it is necessary to provide clear information, supported by scientific evidence. This paper aims to examine the empirical evidence on the social outcomes of the YG program in the countries of Eastern Europe. The present analysis is based on a systematic review, a scientific method is effective for both decision-makers and the scientific community because the data they use and submit to the analysis are classified and evaluated based on scientific criteria, and objectives that allow obtaining a complete, up-to-date, and reliable images. The credibility of the investigated sources is ensured by including in the analysis only the reports published by the European institutions (European Commission, European Council, etc.) regarding YG. Through the analysis of these documents, we found that YG led to an increase in the employability rate of young people, their level of education, and social integration, to the development of self-confidence in their competencies but also in institutions (schools, PES, employers, etc.).

Keywords EU cohesion policy · Youth guarantee

1 Introduction

The YG program, one of the EU's most innovative programs, is the result of the 2013 EU Council Decision (EU Council Recommendation, 2013) which aims to support young people who are considered to be most affected by successive socio-economic crises which member states have faced in the last two decades. Unlike other European programs, YG is differentiated by at least two aspects that give it a higher degree of clarity and predictability: a guarantee and a clearly delimited time period for intervention. (Escudero & Mourelo, 2017). These characteristics are reflected in

G. Neagu (✉)
Research Institute for Quality of Life, Romanian Academy, Calea 13 Septembrie no 13, 050711 Bucharest, Romania
e-mail: gabi.neagu@iccv.ro

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_7

the fact that young people receive a quality offer of employment, further education, retraining, apprenticeship, or an internship within four months of entering or leaving the job or are no longer included in formal education. YG is also characterized by the fact that it addresses a specific age group—15–24 years—and in applying the measures and interventions Member States are free to select their own methods, to build an institutional network to support the implementation of the program, provided that they are adapted to national, regional and local conditions. Although it started at the same time in all EU countries, after almost a decade, the stage of achieving the objectives, the areas in which the YG was applied, the socio-demographic, and economic characteristics of the target group but especially the results differ not only from one country to another, from one region of Europe to another, but also from one dimension of the program to another. In this regard, EU-level reports, and different studies (EC, 2018; Pesquera et al., 2021) point out that better collection and analysis of information on YG results would be needed. Our aim is to analyze the social outcomes of the YG program in the EU member states of Eastern Europe with a focus on two categories of beneficiaries: young people at risk of early school leaving (ESL) and NEETs (young people who are not even in school nor on the labor market)—from the beginning (2014) until now. In this article, we propose a summary of the most relevant results. The first part of the article includes an analysis of the situation of young people in Eastern Europe in terms of ESL and NEETs to understand the context in which the YG program started and took place. The research methodology that we will use to achieve the set goal as well as the results obtained by applying the systematic review method will be found in the second part of the article followed by the third part which includes the discussions. The last part will be dedicated to the conclusions and recommendations.

2 Regional Context

Eastern European countries are characterized by certain aspects that derive from their socio-economic, and historical course: the accession of the East to the EU has occurred recently—the last two decades—in distinct stages and by groups of countries, against a background of socio-economic and institutional development lower. The global socio-economic crises have had a greater negative impact in Eastern Europe because they have overlapped with the national ones generated mainly by the long and difficult transition process from the centralized economy to the market one, from totalitarianism to democracy. European programs of the YG type depend to a large extent on the capacity of national institutions to implement, monitor, etc. Institutional reform, which aims to make them more competent and more transparent, is one of the priority objectives of Eastern European countries. At the same time, the continuous reform can affect the development of YG type programs due to the instability and the frequent changes that it generates. All of this makes the population of young people in Eastern Europe, the target group of YG, to be characterized

Fig. 1 Early leavers from education and training. *Source* Authors' own elaboration. Data available from Eurostat (2022), [edat_lfse_14], Data extracted 08.06.22

by mobility, and diversity but especially by a high level of segmentation from the educational, occupational but also motivational, and aspirational points of view.

2.1 ESL: Main Drivers and Evolution

Eurostat (2022) defines ESL as the situation in which young people aged 18–24 are completed at most a lower secondary education and who are not in further education or training for four weeks. ESL is the result of a combination of factors that go beyond the boundaries of the education system. Young people may leave the education system prematurely due to the quality of education, learning conditions or due to a wrong school orientation but all these situations overlap with belonging to a disadvantaged socio-family and economic environment, an ethnic minority, or living in rural areas, etc. For this category of young people, solutions must be identified that will keep them in school, to restore their confidence in the education system both for themselves and their families. The ESL rate in Eastern Europe between 2014—the year of the program's debut—and 2021 there have been significant variations both from one country to another and within each country from one year to another. The ESL rate varies at the start of YG (2014) from 2.8% in HR to 18.1% in RO (Fig. 1). If we consider the maximum and minimum values of this indicator for 2021, we find that the same countries remain at the base and at the top of the hierarchy, but for both situations, the ESL values decreased: by 0.4% in HR and by 2.7% in RO.

2.2 NEETs: Main Drivers and Evolution

The analysis of the definitions and classifications of NEETs reveals a very high level of heterogeneity: young people who voluntarily choose not to enter the labor market or in the education and training system belong to the same category as young people

Table 1 Young people neither in employment nor in education and training (NEETs rates)

	2014	2015	2016	2017	2018	2019	2020	2021
EU-27 from 2020	15.7	15.2	14.5	13.7	13.1	12.6	13.8	13.1
Bulgaria—BG	24.0	22.2	22.4	18.9	18.1	16.7	18.1	17.6
Czechia—CZ	12.1	11.8	11.1	10.0	9.5	9.8	11.0	10.9
Estonia—EE	13.8	13.2	14.1	11.6	12.1	10.6	11.9	11.2
Croatia—HR	21.8	19.9	19.5	17.9	15.6	14.2	14.6	14.9
Latvia—LV	15.2	13.8	13.3	12.3	11.6	10.3	11.9	12.1
Lithuania—LT	12.9	11.8	10.7	10.2	9.3	10.9	13.0	12.7
Hungary—HU	16.4	15.1	14.1	13.3	12.9	13.2	14.7	11.7
Poland—PL	15.5	14.6	13.8	12.9	12.1	12.0	12.9	13.4
Romania—RO	19.9	20.9	20.2	17.8	17.0	16.8	16.6	20.3
Slovenia—SI	12.9	12.3	10.9	9.3	8.8	8.8	9.2	7.3
Slovakia—SK	18.2	17.2	15.9	16.0	14.6	14.5	15.2	14.2

Source Authors' own elaboration
Data available from: Eurostat (2022), [edat_lfse_20], Data extracted 08.06.22.

who for objective reasons (disease, growth or family care) do not have a job or cannot attend training courses. Three of the Eastern European countries—HU, BG, and RO—are characterized by the highest rates of NEETs from 2014 to 2021 (Table 1). BG and RO are also the least developed countries in Eastern Europe compared to all other EU-27 countries, which is why the effects of socio-economic crises in recent years have had a greater impact and more negative consequences.

Increasing the rate of NEETs individually favors a major risk of marginalization and exclusion from the labor market (Caroleo et al., 2020; Thompson, 2011), impoverishment of human capital (Becker, 1994), and reduced likelihood of future employment in the workplace (Ryan, 2001). At the social level, this leads to a loss of economic productivity and growth (Mascherini, 2019). At the individual level, the status of NEETs over a long period of time has effects on the behaviors and attitudes of young people towards work, and education: low self-esteem, giving up looking for a job, depression, development of anti-social behaviors, etc. Some experts (Robertson, 2019) believe that long periods of educational and/or occupational exclusion make some young people insufficiently prepared in terms of physical and mental health to be the target group for those specific YG interventions that aim to rapid integration in school and/or on the labor market.

3 Methodological Framework

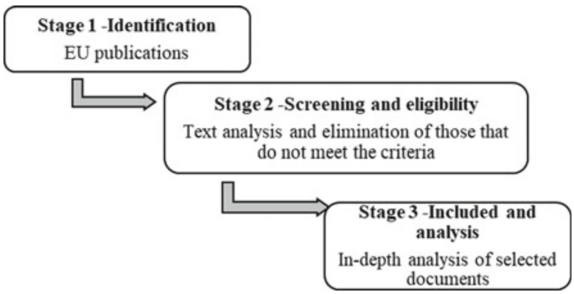
In the last years, the interest in YG implications is demonstrated by both the high number of studies and the diversity of the methods used to be analyzed: comparative evaluation (Ule & Leskošek, 2018), case studies (Tsekoura, 2019), secondary analysis, and public policy analysis (Petrescu et al., 2021), the convergence analysis (Tosun et al., 2019), etc. For this article, we opted for a systematic review method. A systematic literature review is a theoretical approach of which is to reviewing relevant documents in a particular field, documents are selected based on certain criteria of relevant databases (Okoli and Schabram, 2010). The evaluation of the results of the YG program almost a decade after its launch (2013–2022) requires the organization of those conclusions reached so far by differing reports from relevant European institutions and with a high level of credibility. For this purpose, a systematic analysis of the EU's published reports was used.

3.1 *Selection: Criteria and Study*

For the selection of the most relevant documents related to the YG, a series of criteria were specified: (a) the period of time 2014–2022 because the program was launched in 2013; (b) to include the term "youth guarantee" and to refer to the countries of Eastern Europe as a precondition because the main purpose of this study aims to analyze the results of a certain program in a certain regional/territorial area; (c) focus on the dimension of ESL and NEETs or related terms (training, scholarships, etc.); (d) to make references to the type of measures applied, target group, etc. to allow the shaping of a measuring frame; (e) be present in one of the EU publications in order to ensure a high degree of credibility of the information source. The study selection (Fig. 2) was made in three steps.

First, the titles of all retrieved documents were screened for eligibility for the above-mentioned inclusion criteria. Second, the summary or introduction of all initially relevant documents was screened for eligibility by applying uniform criteria.

Fig. 2 Document selection process. *Source* Authors' own elaboration



Finally, the full text of all remaining publications was analyzed. After applying the selection criteria, 87 documents were selected for this review.

3.2 Results

The evaluation of the YG program can be done from various perspectives: evaluating the social profitability and effectiveness of interventions for the various beneficiaries of the program, evaluating the interaction and complementarity of the program with a country's broader social welfare policies, evaluating social outcomes and social and economic impact of YG on individuals and communities, etc. In this article, we will focus on evaluating YG's social outcomes. We focused on programs targeting ELS and NEETs because investments in education are, according to specialists (Becker, 1994), the most profitable both individually and socially. Table 2 shows an overview of the target group and the type of intervention likely to have an impact both on the individual and on the communities they belong to.

In order to evaluate the social outcomes of the YG on the beneficiaries, we grouped the information obtained by analyzing the documents selected into two main categories: target group and type of intervention (Table 2). Within them, there are different

Table 2 Overview of target groups and types of intervention through the YG program

Countries	Target group			Type of intervention
	Type	Age	Specific target group	
BG, LV, CZ	NEETs	15–29	–	Employment and training programs, Vocational programs
BG	NEETs	15–24	Roma population	Training program
RO	NEETs	16–24	NEETs from rural areas and Roma	Creation of an electronic registry for NEETs
HR	ELS	–	–	Career guidance and counseling
SK	ELS	16–24	Roma population	Mentoring program
SI	ELS	15–26	–	Learning program for young adult
PL	ESL	–	–	Training to improve literacy and numeracy
LT	NEETs	16–25	–	Counseling to build up confidence and motivation
EE, LV, LT, PL HU, SK, RO	NEETs	16–29	–	Subsidized employment offers

Source Authors' own elaboration

sub-categories (age of beneficiaries, specific group, etc.) that support the in-depth analysis of the topic. In selecting the categories of analysis, we took into account the recommendations of various guidelines developed at the EU level: e.g. Guidelines to analyze the performance of the national YG (ILO, 2016).

4 Discussions

A synthetic analysis of the main results obtained by applying the systematic review method reveals some specific features both to the two categories of analysis—target group and type of interventions. The beneficiaries of the YG actions aimed at increasing the level of education and/or support for integration into the labor market were young people aged 15–29. Through the YG program, young people over the age of 24, but not older than 30, can be considered a target group if member countries deem it necessary. The analysis of the documents reveals that most Eastern European countries have extended the maximum age range from the target group included in the YG from 24 to 29 years. In all Eastern European countries, more than young ESL, NEETs were more frequently targeted by YG actions in terms of both education and professional integration. This is not only because the NEETs rate is above the EU-27 average but also because in some Eastern European countries (CZ, EE, HU, LT, LV, RO, SI) young NEETs have a low level of education. In this case, the intervention must be carried out in both directions: increasing the level of education and support for integration into the labor market. By comparison, BG, SK, PL, and HR have a high level of education among young NEETs. (Botrić, 2017; Caliendo et al., 2019) requiring a different approach: the emphasis must be moved from the basic needs (a minimum level of education and facilitating professional integration) to those of motivation and achievement: the adequacy of the job offer to the level of education and aspirations, support for entrepreneurship, etc. Even if in the literature of specialty but also in the official documents of the EU, regarding age are included in the NEETs category, young people aged 15–24, some researchers (Maguire & Thompson, 2007) consider this interval to be far too restrictive given the particularities of this stage of life—youth—but also the issues that the general population must face today. The results of the analysis lead to the idea that the need for support is even higher for the age groups 24–29 years and we find that the positive impact of interventions is higher. Thus, in BG, SK, EE, HR, HU, LV, and RO the 25–29 NEET rate decreased by 1–4 pps, and the employment rate increased by more than 1–5 pps. (EC, 2016) following the application of YG. The data from the analysis and the results obtained by some countries by applying YG support the opinion of British researchers: young NEETs also need psychological support, as well as training in basic or professional skills, as well as a quality job offer whatever their age. From social outcomes, this means reducing the unemployment rate and dependence on PES. The fact that in some Eastern European countries, the most vulnerable category of populations in terms of education and employment is the Roma minority (Cace and et al., 2014) is also reflected in the YG program: BG and RO have established that young people belonging to this minority are a priority target group. The interventions addressed to this specific target group are more complex because they must include the young person's family, as well as education, the development of confidence and motivation to learn, counseling for adaptation and professional integration, etc. Another specific group is young Romanian rural NEETs who represent a social problem for RO (Neagu, 2020). In this case, the biggest problem is identifying the beneficiaries because in most countries, for example, RO, these young people tend to be "lost" statistically due to temporary migration, lack of information, etc. Another problem that young people with a high degree of vulnerability raise is that they do not become dependent on YG programs. To avoid this situation the type of intervention applied in each country and for each target group is very important. Also, evaluating the results of YG interventions in each target group supports the successful implementation of the program.

YG allows for 4 major types of intervention: job, apprenticeship, traineeship, and education. Within each type of intervention, different country options can be identified in the selection of the most appropriate programs for the target groups. The analysis of the documents related to YG highlights the fact that they aimed primarily at developing educational and practical skills to help young people with difficulty continue or complete their level of education or integrate into the labor market. It is also found that in many cases the success of YG is due to the development of programs that have helped young people to rediscover and identify, and define their interests, goals, and educational and occupational talents. This type of intervention mainly targeted disadvantaged young people who have experienced long periods of exclusion and educational, social, and professional exclusion and who need a transition period to regain confidence in themselves and in social systems. According to Robertson (2019), these young people must first be psychologically prepared in order to be later supported for their integration into the labor market. For this category of beneficiaries, actions have been carried out within the YG (EC, 2016) to help build confidence and motivation (LT), and to facilitate the experimentation of several types of programs to help them rediscover their competencies. (LV), rediscovering their motivation for learning (SK). Another type of intervention is aimed at young people with potential, who are at a disadvantage because they have not had the opportunity to capitalize on their talents, and skills or because they have not been guided by a directive to put them in place value skills. The most frequent interventions fall into the category of career guidance (HR) and consist of youth entrepreneurship programs, grants for business start-ups, guidance and/or financial support for business plan development, etc. It should be mentioned that even in these situations the measures are combined with another type of support (EC, 2019): the one to help them cope with the stress, and the risks involved in entrepreneurship. Supporting entrepreneurship means not only the integration of young people but also the creation of new jobs for other young people in difficulty. We observe that the dominant type of intervention at the level of a country correlates with the dominant type of target group: in countries (HU, RO, e.g.) where dominant groups are young people with very low levels of education, characterized by long periods of social and professional exclusion, the emphasis is on completion, compensation; in countries (HR or EE) where young people have at least an average level of education and a minimum of experience on the labor market, the emphasis is on development and capitalization of skills. The particularities of the countries but also of the target group are reflected in the visibility of the obtained results: much more visible when the completion of the level of education and integration in work was aimed at and less visible in the case of institutional development, training programs, counseling, etc. but this will have effects in the medium and long term.

5 Conclusions and Further Research

The partial results presented in this article claim that most Eastern European countries have extended the age of the target group of YG and have achieved very good results in the age category over 24 years which means that changes should be taken into account in the age of target group level; each country opted for the implementation of the YG by methods appropriate to the needs of the target group, which contributed to the success of the program; there are differences between Eastern European countries regarding the target groups, the type of intervention, results, etc. but this must be put in context in order to understand objectively what the reasons are and what the effectiveness of YG is especially in terms of social outcomes. Every successful intervention in YG represents a gain for both the beneficiary and the community. In the case of those where the intervention aims to increase the level of trust in social systems, it means a step in breaking the vicious circle of intergenerational, community poverty; in the case of those who need support in order to capitalize on their entrepreneurial skills, it means increasing job offers. The effectiveness and efficiency of programs like YG depend to a very large extent on the existence of reliable information, supported by scientific evidence. Documents published by the official EU website are an important starting point in identifying the need for public policies in a sector or field, and the most appropriate type of intervention. The more scientific evidence there is and the higher the level of trust, the greater the chances that the field, the sector in question, will be supported. This can be seen from the summary presented in this article: the systematic review method supports the identification of YG target groups and their particularities, and the type of intervention chosen in each country, allows comparability between countries on different dimensions of YG. Data collected for this analysis will be analyzed more complexly, in more detail to better understand the situation of the YG program and to contribute important information for decision-makers.

References

Becker, G. S. (1994). Human capital revisited. In Human capital: A theoretical and empirical analysis with special reference to education (3rd edn., pp. 15–28). Chicago, IL: The University of Chicago Press.

Botrić, V. (2017). Youth employment initiative (YEI) in Croatia: in depth-analysis. European Parliament. https://data.europa.eu/doi/10.2861/236129.

Cace, S., Neagu, G., Raţ, C., Ivasiuc, A. (Eds.) (2014). Politici de incluziune a romilor din statele membre ale UE. Bucureşti: Editura Institutul European din România.

Caliendo, M. & Schmidl, R. (2016). Youth unemployment and active labor market policies in Europe. IZA Journal of Labor Policy *5*(1), 1–30. https://doi.org/10.1186/s40173-015-0056-3.

Caliendo, M., Kluve, J., Tübbicke, S., & Stöterau, J. (2019). Study on the youth guarantee in light of changes in the world of work. Part 1, Youth guarantee: intervention models, sustainability and relevance. European Commission, Directorate-General for Employment, Social Affairs and Inclusion, Publications Office. https://data.europa.eu/doi/10.2767/371432.

Caroleo, F. E., Rocca, A., Mazzocchi, P., & Quintano, C. (2020). Being NEET in Europe before and after the economic crisis: an analysis of the micro and macro determinants. Social Indicator Research, *149*(3), 991–1044. https://doi.org/10.1007/s11205-020-02270-6.

European Commission. Employment, Social Affairs & Inclusion. (2018). Data collection for monitoring of Youth Guarantee schemes. https://ec.europa.eu/social/main.jsp?catId=1143&langId=en#YGIF

European Commission, Directorate-General for Employment, Social Affairs and Inclusion, Luminita Anghel, L. (2019). Report on PES implementation of the Youth Guarantee. September 2019, Publications Office. https://data.europa.eu/doi/10.2767/625047.

European Commission, Directorate-General for Employment, Social Affairs, and Inclusion. (2018). Employment and entrepreneurship under the Youth Guarantee: experience from the ground. Publications Office. https://data.europa.eu/doi/10.2767/053196.

European Commission Directorate-General for Employment Social Affairs and Inclusion Directorate B—Employment. (2019). Assessment of the people not in employment, education and training (NEETs) in Bulgaria and policy measures to effectively address their integration. https://op.europa.eu/en/publication-detail/-/publication/ca757f7f-6e48-11ea-b735-01aa75ed71a1/language-en/format-PDF/source-259397305.

Escudero, V., & Mourelo, E. L. (2017). The European Youth Guarantee: A systematic review of its implementation across countries. ILO. Working Paper no. 21. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_572465.pdf.

Eurostat. (2022). Early leavers from education and training. Eurostat. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Early_leavers_from_education_and_training.

ILO. (2016). Guidelines to analyze the performance of the national Youth Guarantee. https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/publication/wcms_544348.pdf.

Maguire, S., & Thompson, J. (2007). Young people not in education, employment, or training (NEET) where is Government policy taking us now? *Youth and Policy, 8*(3), 5–18.

Mascherini, M. (2019). Origins and future of the concept of NEETs in the European policy agenda. In Youth Labor in Transition. Oxford University Press (pp. 503–529).

Neagu, G. (2020). Rural NEETs in Romania. www.rnyobservatory.eu.

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. Working papers on information systems *10*(26), 1–51. https://edisciplinas.usp.br/pluginfile.php/4126343/mod_resource/content/1/systematic%20literature%20reviews%20Okoli%2C%20Schabram%202010%20Sprouts.pdf.

Pesquera Alonso et al., 2021 Pesquera Alonso, C., Muñoz, S. P., & Iniesta Martínez, A. (2021). Youth Guarantee: Looking for explanations. Sustainability *13*, 5561. https://doi.org/10.3390/su13105561.

Petrescu, C., Neguț, A., & Mihalache, F. (2021). Implementation of the Youth Guarantee programme in Romania. Calitatea Vieții *32*(4), 449–467. https://doi.org/10.46841/RCV.2021.04.07.

Council recommendation of 22 April 2013 on establishing a Youth Guarantee, 2013/C 120/01. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32013H0426%2801%29.
Robertson, P. J. (2019). The casualties of transition: The health impact of NEET status and some approaches to managing it. *British Journal of Guidance & Counselling, 47*(3), 390–402. https://doi.org/10.1080/03069885.2018.1455168
Ryan, P. (2001). The school-to-work transition: A cross-national perspective. *Journal of Economic Literature, 39*(1), 34–92.
Thompson, T. (2011). Individualisation and social exclusion: The case of young people not in education, employment or training. Oxford Review of Education *37*(6), 785–802. https://doi.org/10.1080/03054985.2011.636507.
Tosun, J., Treib, O., & De Francesco, F. (2019). The impact of the European Youth Guarantee on active labour market policies: A convergence analysis. *International Journal of Social Welfare, 28*(4), 358–368.
Tsekoura, M. (2019). A case study analysis of the Youth Guarantee Scheme implementation in Italy (Lombardy Region). *Young, 27*(5), 486–501. https://doi.org/10.1177/1103308818817891
Ule, M., & Leskošek, V. (2018). Transition from education to employment: Comparative assessment of Youth Guarantee policies in Slovenia, the Czech Republic, and Latvia. In: Malo, M., & Moreno Mínguez, A. (Eds.) European youth labour markets. Cham: Springer. https://doi.org/10.1007/978-3-319-68222-8_8.

Methodologies for Supporting the Selection of Projects

Smart and Sustainable Scheduling of Charging Events for Electric Buses

Padraigh Jarvis, Laura Climent, and Alejandro Arbelaez

Abstract The Irish transportation sector currently accounts for more than 30% of the energy related CO_2 emissions of the country. Therefore, in order to reach the sustainable goals, the Irish government is working on multiple incentives to promote Electric Vehicles (EV) and infrastructure to decarbonize the sector, e.g., free domestic charging points, tool reductions, and the implementation of electric Buses (eBuses) in the medium to long term. In particular, eBuses operate with rechargeable batteries with a capacity to store approximately 300 kWh (and up to 600 kWh), equivalent to around 29.9 L of diesel, while reaching approx. 200 km. In order to ensure a proper transition from regular diesel buses to eBuses, charging times must be coordinated to ensure each bus has adequate energy to complete their operational route. In this work, we present a framework for an efficient management of renewable energies to charge a fleet of eBuses without perturbing the quality of service. Our framework starts by building a deep learning model for wind power forecasting to predict clean energy time windows, i.e., periods of time when the production of clean energy exceeds the demand of the country. Then, the optimization phase schedules charging events to reduce the use of non-clean energy to recharge eBuses while passengers are embarking or disembarking. The proposed framework is capable of overcoming the unstable and chaotic nature of wind power generation to operate the fleet without perturbing the quality of service. As expected, the size of the batteries does have a positive impact on the percentage of clean energy required to operate large fleets of eBuses. Methods developed in this paper help to mitigate potentially inaccuracies derived the prediction models. Our extensive empirical validation with real instances from Ireland suggests that our solutions can significantly reduce non-clean energy consumed on large datasets.

Keywords Optimization · Scheduling · Electric buses

P. Jarvis (✉)
School of Computer Science and Information Technology, University College Cork, Cork, Ireland
e-mail: p.jarvis@cs.ucc.ie

L. Climent · A. Arbelaez
Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_8

1 Introduction

The Vehicle Routing Problem (VRP) is a well-known combinatorial optimization problem with applications ranging from logistics to planning and scheduling. This problem involves the creation of optimal routes (e.g., minimizing the traveled distance or the required time to complete certain tasks). These routes might represent supply chains where vehicles deliver goods from a set of depots to customers (Laporte & Nobert, 1987). Research into the usage of EVs has spawned a variation of the VRP called the Electric Vehicle Routing Problem (EVRP). EVRP differs from traditional VRPs as the range of EVs is considerably shorter compared to traditional combustion vehicles. As pointed out above, the range of EVs varies depending on multiple factors, e.g., battery size, average speed, and ambient temperature. Furthermore, some form of charging must occur to complete the daily operations of the vehicles (in particular, for variations of the problem with pick-ups and deliveries (Olgun et al., 2021)). The EVRP focus mainly on minimizing the total cost of routing strategies (Lin et al., 2016) and the placement of charging stations to minimize or even negate detours needed to charge (Funke et al., 2015, 2016).

The Vehicle Routing Problem with Time Windows (VRPTW) is a popular variation of the traditional VRP, where vehicles must visit a set of customers within certain predefined time periods (e.g., outlined by the customers or local governments). This adds additional complexity to VRPs as a vehicle arriving early to a destination might be required to wait, and a vehicle arriving late may invalidate the solution (Desrochers et al., 1992). This has also spawned additional variations such as Time Window Assignment Vehicle Routing Problem (TWAVRP) focusing on assigning time windows to deliveries before the demand is known (Spliet & Gabor, 2015).

Variations of the VRPTW for EVs have also received significant attention recently. The Electric Vehicle Routing Problem with Time Windows (EVRPTW) aims at creating optimal routes as the traditional VRPTW, however the additional constraints of battery capacity and location on recharging stations are also taken into consideration (Schneider et al., 2014). Another line of work considers the charging location problem of EVs. One notable work in this area focuses on the transition to eBuses and the authors proposed a Mixed Integer Programming (MIP) model to identify suitable locations of fast charging units to maintain the current level of service, i.e., same routes and similar timetables (Arbelaez & Climent, 2020).

With the increase in research around EV there has also been a rising interest in using renewable energy to charge EVs. (Zhang et al., 2013) proposed the use of locally generated renewable energy to supplement the requirements of acquiring energy from the national grid. However, when creating a bus operation schedule information such as available renewable energy is needed ahead of time. Predict and Optimize (Elmachtoub & Grigas, 2021) is a relatively new paradigm which focuses on combining predictions and combinatorial optimization. The paradigm involves two stages: the first one involves training a model (e.g., a supervised learning or a time series model) to predict critical variables of the optimization problem. The second stage then uses these predicted values to solve an optimization problem e.g., weights in the weighted knapsack problem (Mandi et al., 2020) or scheduling of combinatorial problems with uncertain duration times (Duque et al., 2018). In this paper, we also use this two-stage paradigm. We start with a time series model to estimate surpluses of wind power in the national grid and then optimize the scheduling of charging events based on the predictions. A time series is a collection of consecutive measurements of powers in kWh recorded in equal intervals (15 min in this paper). The accuracy of the time series methods varies considerably with different forecasting horizons (number of future observations). In this paper we focus on medium-term horizons, i.e., the forecasting period ranges from 6 h to 1 day ahead. A 6, 12, 18, and 24 h ahead forecasting horizon will predict respectively a total of 24 (4 per hour $\times$ 6), 48, 72, and 96 observations. (Shobana Devi et al., 2021) outlines alternative models for other forecasting horizons, i.e., very short-range (a few seconds to 30 min ahead), short-range (30 min–6 h ahead), and long-term range forecasting (1 day to a week ahead).

Long Short-Term Memory (LSTM) is a popular deep learning architecture proven to be effective at energy forecasting (Lim & Zohren, 2021). Such models can be trained to make multi-step ahead predictions, where a variable n controls the number of future time-step predictions (Sangiorgio & Dercole, 2020).

2 Predict then Optimize Framework

The predict and optimize framework aims at guiding the optimization solver to tackle complex problems. In particular, we use a LSTM model to predict how much excess wind energy is available at any time period. This information is then passed to a MIP solver to identify suitable schedules to operate the fleet of eBuses while satisfying certain properties of the transportation system (Arbelaez & Climent, 2020).

2.1 Prediction Models

As pointed out above, we create four LSTM models for predicting excess wind energy 6, 12, 18, and 24 h in advance. Furthermore, we populate our models with historical data from the Irish nation grid in 15 min intervals and populate the training dataset with data from August 2013 to October 2021, and test dataset with data from November 2021 to January 2022 (demand and wind generated power dataset is available at http://smartgriddashboard.eirgrid.com/). The months of November 2021 to January 2022 where selected due to the increased about of wind power generated in winter months, therefore the ability of the LSTM model to predict excess in clean energy can be more accurately determined. Furthermore, we reserve 33% of the training dataset as a validation dataset. The demand and wind generated power datasets are aggregated into one dataset which represents the excess of clean energy

Table 1 MAPE and RMSE for LSTM models with different forecasts times in hours

Forecast time	MAPE	RMSE
6	113.0	313.56
12	186.3	801.75
18	251.1	1048.98
24	292.2	1106.39

Source Authors' own elaboration

at any time. However, at the moment Ireland's national grid does not supply enough wind power to cover the demand. Therefore, we scaled the amount of wind power by a 1.4 factor to simulate a transition to eBuses with enough power to satisfy the current demand. This is in line with the estimations for Ireland's growth in wind generated power by 2026/2027 (Department of Communications, Climate Action & Environment, 2019, p. 40). Therefore, we use a univariate dataset consisting of values between −5064.2 (representing a clean energy deficit of 5064.2) and 1005.8 (representing a clean energy surplus of 1005.8). The LSTM models are then trained on this data with a loss function of Root Mean Squared Error (RMSE) and using Adam optimizer (Kingma & Ba, 2014). We make the data stationary by applying a difference operation on each subsequent value and normalized to a range between −1 and 1 with minMax normalization.

Table 1 shows the Mean Absolute Percentage Error (MAPE) and RMSE of the LSTM models based on the results of the test dataset. We remark that these results are consistent with the literature, i.e., shorter prediction horizons produce more accurate results (Shobana Devi et al., 2021). This increase in error as the horizons become larger can be explained by the fact wind power is harder to predict due to the unstable and chaotic nature of wind power derived from multiple factors, e.g., wind speed, air density, wind turbines, etc.

2.2 *Optimization Model*

We extended the work of (Arbelaez & Climent, 2020) with additional constraints to determine whether the charge times of eBuses overlaps with times where there is an excess of clean energy. Furthermore, our solver aims at reducing CO_2 emissions, and therefore, we minimize the total amount of non-clean energy used to operate the system. In our simulations, we assume that the eBuses travel at a constant speed of 35 km/h and a charge rate of 10 kWh per minute.

We also assume that the discharge rate of the batteries is 1 kWh per km.

We test multiple battery capacities for the bus fleets, these include 120, 180, and 240 kWh. We assert that the battery capacity must not fall below 12 kWh and only allow buses to charge up to 80% of their maximum capacity in a single charge. We also simulate a degree of overnight charging before the buses operational day begins. To represent this, we assume the buses start with a capacity of 30 kWh regardless of the selected battery capacity. It is assumed the placement of charging stations is a separate problem to the one covered in this paper (Loaiza Quintana et al., 2022), to this end the location of charging stations is passed as an argument to the optimization code. For this paper, we assume that charging stations are placed every × km on each bus route. Alternative placement methods involving cost functions will be explored in future works.

We evaluate our framework on three Irish cities, i.e., Cork, Limerick, and Galway. The bus system in Cork includes 11 bus routes operated with 81 buses and 578 stations; the network in Galway includes 6 bus routes operated with 24 buses and 288 stations; and the network in Limerick includes 6 bus routes operated with 23 buses and 253 routes (GPS location of bus stations and timetables are available at https://www.buseireann.ie). Furthermore, we assume two charging infrastructures. The first one, Inf-A, assumes that there is a charging unit every 12 km which results in 53 charging stations for Cork, 13 for Limerick, and 19 for Galway. The second one, Inf-B assumes that there is charging unit every 15 km which results in 43 charging stations for Cork, 12 for Limerick, and 11 for Galway.

We also assert a maximum deviation time for the newly created schedule, meaning the arrival times in the new schedules can be at most Δ different from the original schedule where Δ is an amount of time in minutes. For this paper we explore two values for Δ, 5 and 10 min.

3 Experiments and Results

Prediction models are used to predict wind energy excess for the two-week period of 14th to the 27th of February 2022. The month of February is chosen as it features a high amount of wind generated power, as a result there will be enough excess power to evaluate the performance of our framework. For the optimization model we assume three different scenarios regarding clean energy information. The first $|\Gamma|=0$ assumes the optimization model has no knowledge of clean energy information. The second uses the information generated by the prediction models previously outlined. Finally, we examine the ideal scenario, where we have perfect predictions (i.e., the actual historical values for excess wind energy).

Figure 1 shows the amount of modified wind energy vs the demand of the electrical grid. Of note there are a number of days where the amount of modified wind energy does not exceed system demand at any point (i.e., the 17th, 18th, 21st, and 24th). As a result, experiments which use wind data from these days will produce poor results as there is no clean energy available. On the contrary, the 26th features a very high amount of wind energy throughout the day, as a result any charges which take place on this day would use clean energy. For the empirical analysis of our experiments, we removed the results from the previously mentioned 5 days as they would represent outliers in the amount of wind energy available. Such outliers would not provide any insights into the performance of our framework, as the framework aims to reduce

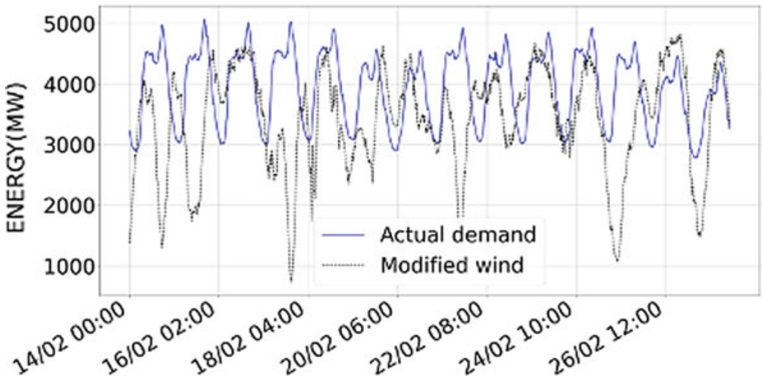


Fig. 1 Energy demand versus 1.4 times modified wind energy for 14/02/2022–27/02/2022. *Source* Authors' own elaboration

the total amount of non-clean energy used and therefore requires clean energy to operate.

Table 2 shows the results for the experiments using Inf-A and a deviation time of five minutes. As expected, when using predictions from our prediction models there is a notable decrease in the amount of non-clean energy used when compared to a naïve scenario with no knowledge of clean energy information. Therefore, while our LSTM models are not fully accurate there is a notable benefit in the integration of the learning component in our schedules. Also of note is that increasing battery capacity in the ideal scenario may not reduce the amount of non-clean energy consumed. This is because all of the available clean energy is already being consumed when capacity is set to 120. This is the case on the smaller datasets of Limerick and Galway; however, we see for the Cork dataset that increasing the capacity does reduce the non-clean energy consumed. Sometimes when using predictions from our LSTM models larger battery capacities consume more non-clean energy compared to smaller capacities, we attribute this to mispredictions in our LSTM models. Larger battery capacities can consume more energy in a single charge. As a result, any false positives in our prediction model (i.e., we predict there is a clean energy excess when there is actual a deficit) could result in the scheduling of a charge using non-clean energy.

Figure 2 shows a comparison between the average across all capacities for each scenario and each city. Here we see that the average difference between the three scenarios heavily demands on the dataset used. For example, smaller datasets like Limerick show minor differences between the three scenarios. However, the larger Cork dataset shows more significant difference. We attribute this to the higher energy requirements of the Cork dataset in addition to the longer operational times of the bus system. The Cork bus system begins operation earlier then both the Galway and Limerick datasets, and finishes operational routes later, as a result the Cork data-set is able to make use of any excess clean energy in the early morning and late night.

Experiments conducted using $\Delta = 10$ min found that the solutions in the ideal scenario only improve by 0.685% on average. It should also be noted that experiments

Table 2 Non-clean energy used (in kWh) for Inf-A and $\Delta = 5$ min

City	Capacities	$\|\Gamma\| = 0$	Predicted	Ideal
Limerick	120	3338.7	3305.1	3108.5
	180	3347.8	3318.5	3108.5
	240	3337.4	3317.2	3108.5
Galway	120	4279.9	4139.3	3850.3
	180	4282.9	4184.4	3850.3
	240	4286.1	4151.0	3850.3
Cork	120	11,458.5	10,970.1	10,302.1
	180	11,446.3	10,987.4	10,180.2
	240	11,423.2	10,939.5	10,148.4

Source Authors' own elaboration

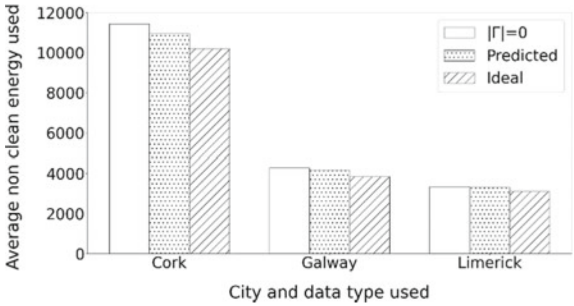


Fig. 2 Average scenario performance per city. *Source* Authors' own elaboration

using $\Delta = 10$ min took 27.39% longer to complete compared to $\Delta = 5$ min. Results for experiments using Inf-B showed an increase in the amount of non-clean energy used on average. However, it should be noted, for the Cork dataset and capacity 240 solutions where only 0.93% worse compared to experiments using Inf-A. As the reader recalls Inf-A for the Cork dataset features 53 charging stations, while Inf-B has 43. This suggests the relationship between number of charging stations present and solution quality are not directly proportional.

4 Conclusions and Further Research

In this work, we use a deep learning model for wind power forecasting to estimate the availability of clean energy in a day, we then integrate the output into an optimization model to schedule charging events. Experimentation results with actual data from the Irish national grid and a major bus operator in Ireland suggest our models can make a notable reduction in the non-clean energy consumed compared to a naïve optimizer. While our predictions do not generate solutions as high quality as the ideal scenario, a significant reduction in non-clean energy consumed can be observed on larger datasets. Therefore, the results of the evaluation confirm the high-quality performance of the proposed approach. In the future, we plan to extend our framework with Bus-to-grid technology to help the national grid by returning energy when needed (i.e., during peak hours). We also plan to investigate the performance of our proposed framework with the charging infrastructure placement problem.

Acknowledgements This work received funding from the Sustainable Energy Authority of Ireland (SEAI) Research, Development & Demonstration (RDD) 2019 programme under the grant number 19/RDD/519. The authors would like to thank the anonymous reviewers for their comments and suggestions which helped to improve the paper.

References

Arbelaez, A., & Climent, L. (2020). Transition to ebuses with minimal timetable disruptions. In *Thirteenth Annual Symposium on Combinatorial Search*, Vienna, Austria, May 26–28.

Department of Communications, Climate Action & Environment. (2019). *National Climate & Energy Plan 2021–2030*. Retrieved July 19, 2022, from https://assets.gov.ie/94442/f3e50986-9fde-4d34-aa35-319af3bfac0c.pdf

Desrochers, M., Desrosiers, J., & Solomon, M. (1992). A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research, 40*(2), 342–354.

Duque, R., Arbelaez, A., & Díaz, J. F. (2018). Online over time processing of combinatorial problems. *Constraints an International Journal, 23*(3), 310–334.

Elmachtoub, A. N., & Grigas, P. (2021). Smart "predict, then optimize." *Management Science, 68*(1), 9–26.

Funke, S., Nusser, A., & Storandt, S. (2016). Placement of loading stations for electric vehicles: Allowing small detours. In *Twenty-Sixth International Conference on Automated Planning and Scheduling*, London, UK, June 12–17 (Vol. 26, No. 1, pp. 131–139).

Funke, S., Nusser, A., & Storandt, S. (2015). Placement of loading stations for electric vehicles: No detours necessary! *Journal of Artificial Intelligence Research, 53*, 633–658.

Kingma, D. P., Ba, J. (2014) Adam: A method for stochastic optimization. arXiv preprint arXiv: 14126980

Laporte, G., & Nobert, Y. (1987). Exact algorithms for the vehicle routing problem. In: S. Martello, G. Laporte, M. Minoux & C. Ribeiro (Eds.), *North-Holland Mathematics Studies* (Vol. 132, pp. 147–184).

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A, 379*(2194).

Lin, J., Zhou, W., & Wolfson, O. (2016). Electric vehicle routing problem. *Transportation Research Procedia, 12*, 508–521.

Loaiza Quintana, C., Climent, L., & Arbelaez, A. (2022). Iterated local search for the ebuses charging location problem. In *Seventeenth International Conference on Parallel Problem Solving from Nature*, Dortmund, Germany, September 10–14 (To Appear).

Mandi, J., Stuckey, P. J., & Guns, T. (2020). Smart predict-and-optimize for hard combinatorial optimization problems. In *Thirty-Fourth AAAI Conference on Artificial Intelligence*, New York, USA, February 7–12 (Vol. 34, No. 2, pp. 1603–1610).

Olgun, B., Koç, Ç., & Altiparmak, F. (2021). A hyper heuristic for the green vehicle routing problem with simultaneous pickup and delivery. *Computers & Industrial Engineering, 153.*

Sangiorgio, M., & Dercole, F. (2020). Robustness of lstm neural networks for multi-step forecasting of chaotic time series. *Chaos, Solitons & Fractals, 139.*

Schneider, M., Stenger, A., Goeke, D. (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science, 48*(4), 500–520.

Shobana Devi, A., Maragatham, G., Boopathi, K., & Lavanya, S. R. (2021). Long-term wind speed forecasting—A review. *Artificial Intelligence Techniques for Advanced Computing Applications, 130*, 79–99.

Spliet, R., & Gabor, A. F. (2015). The time window assignment vehicle routing problem. *Transportation Science, 49*(4), 721–731.

Zhang, T., Chen, W., Han, Z., et al. (2013). Charging scheduling of electric vehicles with local renewable energy under uncertain electric vehicle arrival and grid power price. *IEEE Transactions on Vehicular Technology, 63*(6), 2600–2612.

Multiple Impacts of Energy Efficiency Technologies in Portugal

Marcos Tenente, Carla Henriques, Álvaro Gomes, Patrícia Pereira da Silva, and António Trigo

Abstract Portuguese programs aimed at fostering Energy Efficiency (EE) measures often rely on cost–benefit approaches only considering the use phase and neglecting other potential impacts generated. Therefore, this work suggests a novel methodological framework by combining Hybrid Input–Output Lifecycle Analysis (HIO-LCA) with the Portuguese seasonal method for computing the households' energy needs. A holistic assessment of the energy, economic, environmental, and social impacts connected with the adoption of EE solutions is conducted aimed at supporting decision-makers (DMs) in the design of suitable funding policies. For this purpose, 109,553 EE packages have been created by combining distinct thermal insulation options for roofs and façades, with the replacement of windows, also considering the use of space heating and cooling and domestic heating water systems. The findings indicate that it is possible to confirm that various energy efficiency packages can be used to achieve the best performance for most of the impacts considered. Specifically, savings-to-investment ratio (SIR), Greenhouse gases (GHG), and energy payback times (GPBT and EPBT) present the best performances for packages that exclusively employ extruded polystyrene (XPS) for roof insulation (packages 151 and 265). However, considering the remaining impacts created by the investment in energy

M. Tenente (✉) · C. Henriques · Á. Gomes · P. P. da Silva
Institute for Systems Engineering and Computers at Coimbra (INESC Coimbra), University of Coimbra, Electrical and Computer Engineering Department (DEEC), Rua Sílvio Lima, Polo II, 3030-290 Coimbra, Portugal
e-mail: marcos.tenente@inescc.pt

C. Henriques · A. Trigo
Polytechnic of Coimbra, Coimbra Business School, 3045-601 Coimbra, Portugal

Á. Gomes
University of Coimbra, Electrical and Computer Engineering Department (DEEC), Polo II, Rua Sílvio Lima, 3030-290 Coimbra, Portugal

C. Henriques · P. P. da Silva
University of Coimbra, Centre for Business and Economics Research (CeBER), Faculty of Economics (FEUC), Av Dias da Silva 165, 3004-512 Coimbra, Portugal

A. Trigo
Algoritmi Research Center, University of Minho, 4800-058 Guimarães, Portugal

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_9

efficiency measures, their best performances are obtained when roof and façades insulation is combined with the use of space heating and cooling and DHW systems to replace the existing equipment. If biomass is assumed to be carbon–neutral, solution 18,254 yields the greatest reduction in GHG emissions. Given these trade-offs, it is evident that multiobjective optimization methods employing the impacts and benefits assessed are crucial for helping DMs design future EE programs following their preferences.

Keywords Energy efficiency · Hybrid input–output lifecycle analysis · Multiple benefits

1 Introduction

Considering the current energy standards, 75% of the European Union (EU) buildings are inefficient and more than 85% of these will still be operating in 2050. Additionally, the building sector accounts for about 40% of energy consumption and 36% of GHG emissions in the EU. Therefore, in the scope of EU long-term strategy of carbon neutrality by 2050, was recognized the need of accelerating the renovation rate of the European buildings to reach a carbon–neutral competitive economy and promoting growth and job creation (European Commission, 2019; 2021).

In Portugal, the residential building stock presents a similar behavior being responsible for more than 30% of final energy consumption, which increased by 1.6% during the period 2014–2019 (Energy Observatory, DGEG & ADENE, 2021). Also, about 66% of the Portuguese buildings were built before 1990, the year when EE requirements were introduced for new buildings and approximately one-third of the building stock built before 2012, reveals repair needs on the roof and exterior façades, leading to low energy performance levels, thus contributing to energy poverty, energy consumption and emissions generation (INE, 2012). To address the urgency of accelerating the buildings' renovation, Portugal has been deploying several programs aimed at promoting EE growth in this sector of which we can highlight the energy efficiency Fund, the support program for more sustainable buildings and the energy consumption efficiency promotion plan (in Portuguese—PPEC) (Presidency of the Council of Ministers, 2013, 2020, 2021). However, the evaluation of the EE measures to be funded is usually grounded on cost–benefit bases, mainly accounting for energy and emission savings during the operation phase, thus neglecting other potential benefits and impacts connected with all lifecycle (LC) phases of the measures selected. An example is that almost all the measures dedicated to promoting the EE in the residential sector, that are part of the Portuguese long-term strategy for the renovation of buildings, rely only on the assessment of costs and calculation of energy and emission savings during their operation phase (Presidency of the Council of Ministers, 2021). Nevertheless, the right assessment of the energy, economic, environmental, and social improvements of investing in EE should allow DMs to make well-grounded decisions when it comes to the choice of which EE measures should be funded in the residential sector. According to Reuter et al. (2020), a broader understanding of the attainable EE multiple benefits is necessary to facilitate the promotion of EE policies. This idea is in line with the European Renovation Wave establishing the need of using a better definition of the best criteria to be considered in the assessment of the energy-related savings in the new funding instruments. In this manner, one the objectives of the Recovery and Resilience Plan (RRP) is to improve the EE in residential buildings aiming at achieving the reduction of energy consumption and GHG emissions, the reduction of energy poverty, the improvement of indoor comfort and air quality and the creation of employment (European Commission, 2020; Ministry of Planning, 2021; Presidency of the Council of Ministers, 2021). In addition, the investment in EE measures may produce benefits other than energy savings and emissions reductions, like poverty alleviation, industrial productivity and competitiveness, energy security, job creation, energy prices moderation and health and well-being related benefits (Ryan and Campbell, 2012).

With the foregoing in mind, this work proposes a novel holistic approach that integrates a Hybrid Input–Output Lifecycle Analysis (HIO-LCA) framework with methods for calculating the energy performance of buildings (a seasonal method employed by the Portuguese building energy certification system), to evaluate the energy, economic, environmental, and social impacts and benefits of investing in EE solutions in the Portuguese residential sector generated throughout their LC assessment, thus helping DMs in the design of suitable EE funding policies.

2 Literature Review

As mentioned before, generally the operation is the only LC phase that is usually accounted for in the design of EE funding programs. However, when the nearly zero-energy buildings strategy is considered, the assessment of other lifecycle phases becomes even more important. Hence, other avenues of research addressing a broader range of impacts are required to support DMs in the design of suitable EE policies. In this context, the economic input–output LCA (EIO-LCA) makes it possible to assess the direct and indirect impacts of the entire economy connected to the production of a product or the provision of a service, avoiding the time-consuming and truncation problems inherent to the LCA approach. However, the EIO-LCA methodology is not free of limitations, as it can suffer from aggregation problems (Hendrickson et al., 1997; Suh, 2006; Crawford, 2009; Säynäjoki et al., 2017). In this context, an HIO-LCA framework should thus be used. This top-down approach pursues the simplification of LCA, extending conventional Input–Output (IO) matrices with environmental, energy, social or economic impacts, accounting for the transactions of all activity sectors/ industries, implying that the boundary of the analysis becomes very broad and inclusive, and the circularity effects are also included (Hendrickson et al., 1998, 2006; Bilec et al., 2006; Strømman et al., 2009; De Carvalho et al. 2016; Singh et al., 2018a, b). Hybrid methodologies have been used in distinct contexts. For example, to assess the employment impacts of renewable energy technologies

(Oliveira et al., 2014; Henriques et al., 2017), to assess the energy consumption and carbon emissions of a residential building during its lifetime (Zhan et al., 2018), or to compute of the embodied and operational energy of residential buildings in Lebanon (Stephan & Stephan, 2014). Nevertheless, the application of this sort of approach in the context of EE actions is not abundant, with only a few studies found in the literature. In this context can be mentioned the application of an energy and environmental extended EIO-LCA model to assess the benefits arising from the tax deduction for energy retrofit actions in the Italian building stock by Cellura et al. (2013), the assessment of the energy, economic, environmental, and social impacts of fostering the investment in electric energy-efficient appliances in India made by Singh et al. (2018a, b), which were later on combined with multiobjective interval portfolio theory in Singh et al. (2019) to support public DMs on the design of EE investment programs regarding different investment strategies. In the field of EE in the Portuguese residential sector, different types of studies have been conducted over the past few years by Asadi et al. (2012), Oliveira et al. (2014), Tadeu et al. (2018) and Henriques et al. (2020), although the impacts, the technologies and/or LC phases considered present some gaps that are intended to be fulfilled with this work.

This paper is organized as follows: in the next section, we describe the methodological framework proposed. Subsequently, some illustrative results are discussed. Finally, some conclusions are drawn, and future work developments are suggested.

3 Methodology

The analysis starts with the identification of retrofit technologies generally employed in the Portuguese residential sector—also known as "business as usual". Subsequently, the corresponding best EE available technologies were chosen by analyzing several existing funding schemes (i.e., PPEC and LTRS PT). The manufacture, packaging, installation, and maintenance (MPIM) phases of the selected measures are then evaluated through the HIO-LCA approach, which combines Portuguese Supply and Use Tables (SUT) for the year 2017 at basic prices with impact data (INE, 2017, 2019; OECD, 2017; Oliveira et al., 2014). To calculate the multiplier effects of each activity or component (the matrices of direct and indirect coefficients) for the chosen energy, economic, environmental, and social indicators, the total output of each relevant activity or component of the technologies under analysis is linked to the corresponding product using the adjusted rectangular IO table. The retrofitting strategies considered in this case study involve the application of six types of insulation systems to the roofs and façades, with five different thicknesses, the replacement of the single-glazed aluminum frame windows with double-glazed aluminum or PVC frame windows combined with ten types of space heating and cooling and DHW appliances (see Table 1). The thickness of the insulation measures considered in this study did not exceed 120 mm, since there is a lack of available data for higher thicknesses and because this value meets and even exceeds the minimum energy requirements set out by the Energy Performance Regulation of Residential Buildings (Ministry of Economy & Employment, 2013).

Table 1 EE measures

Façade insulation		Roof insulation		Windows		Systems					
Type	Thickness (m)	Type	Thickness (m)	Frame	Glass	DHW + Heating	Heating + Cooling	Heating	Cooling	DHW	Backup
EPS	0.04 0.06 0.08 0.10 0.12	EPS	0.04 0.06 0.08 0.10 0.12	Aluminium	double	Conventional boiler	Air-source heat pump	Electric heater	Portable AC	Gas water heater	Solar collector
ICB		ICB									
Rock wool		Rock wool		PVC		Condensing boiler					
Glass wool		Glass wool								DHW Heat Pump	
PUR		PUR		Aluminium	Simple	Biomass boiler			No cooling system		No solar collector
XPS		XPS								Electric water heater	
No insulation	N/A	No insulation	N/A			DHW + Heating heat pump					

Source Authors' own elaboration

Then, using the seasonal method employed by the Portuguese building energy certification system (Decree-Law 101-D/2020), the annual households' energy requirements for space heating and cooling and DHW are calculated, before and after the interventions on the building envelope. Those needs will be later used to assess the impacts linked to the operation phase of the building. The impacts obtained with this approach were the total primary energy savings (TPES) and the energy payback time (EPBT)—as energy-related impacts; employment and gross value added (GVA) produced during the MPIM phases, savings to investment ratio (SIR) and net present value (NPV)—a proxy of the economic impacts; greenhouse gases (GHG) savings and GHG Payback Time (GPBT)—as environmental impacts and impact on the household budget and reduction of premature deaths—as social impacts. The schematic representation of the methodological approach's implementation proposed is presented in Fig. 1. To simulate the impacts of the EE measures adopted, a single dwelling T2, located in Coimbra, built between 1961 and 1991, was chosen as a reference scenario. Regarding the characterization of the constructive solutions, this building does not have insulation on the roof and façades and single-glazed aluminum frame windows are applied in its envelopment, also uses electric heaters for space heating and gas-fired water heater for DHW, while space cooling is guaranteed by natural ventilation.

The remaining characteristics of the building can be found in Pinto and Fragoso (2018) and Tenente et al. (2021).

In this work, the combined implementation of EE efficiency measures is preferred over the application of single ones, for the sake of the maximization of the energy savings of a building as well as the minimization of the costs linked to the installation phase. For example, if insulation is applied to the roofs and façades the energy savings are not the same as the sum of both individual measures and the operational costs would be significantly reduced, since the work could be done sequentially using just a part of the resources. With the foregoing in mind, the construction of different EE packages raises the problem of finding a common lifespan for each technology considered to compute the NPV, SIR, GPBT, and EPBT, for the reason that the

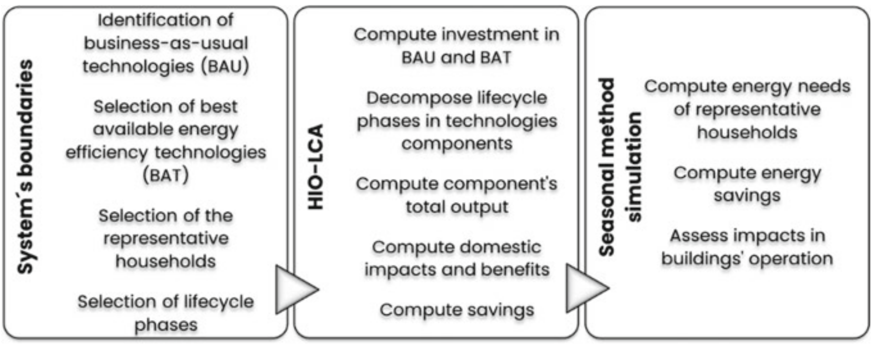


Fig. 1 Schematic representation of the methodological approach. *Source* Authors' own elaboration

lifespan of the insulation measures (50 years) and that of the space heating and cooling and DHW systems (12 to 25 years) do not match, thus being impractical to extend the analysis period to the full life of the efficiency resources being analyzed. On the other hand, if the lifespan of the building was considered, some of the technologies would have to be replaced before, leading to bias in the study. Therefore, the alternative to computing the NPV and SIR is to depreciate the costs of the technologies over their lifespan and to compute the GPBT and EPBT by annualizing the embodied energy (Woolf et al., 2017). It is important to note that the combination of individual EE measures considered led to the construction of new 109,553 packages which will be later compared with our reference scenario in the results section.

Rectangular Input–Output model

Originally developed by Wassily Leontief, IO analysis allows to compute the embodiments of production factors (e.g. labor and energy) and pollutants (e.g. CO2 emissions and waste), per unit of final consumption of commodities, by obtaining the total factor multipliers using IO tables that can be obtainable in diverse structures, according to three main criteria (Miller & Blair, 2009; Sargento et al., 2011): (1) symmetric or rectangular format; (2) total or domestic-use flows; (3) valuation prices (basic prices or purchasers' prices). Rectangular tables at purchaser's prices (in particular, the SUT framework) were firstly introduced by the European System of Accounts (ESA) in 1995, having the ability to consider both the primary and secondary commodities of each industrial sector (Horowitz & Planting, 2006). The Supply/Make matrix is of industry-by-commodity type, giving information on the industrial production of commodities. On the other hand, the Use matrix is of commodity-by-industry type, providing information on the commodities consumed by industries and final users. This format is called rectangular, because the number of commodities included in the model may be higher than the number of industries (Miller & Blair, 2009). Since the SUT framework requires either the consideration of industry or product technology assumption, this work used the industry technology assumption due to the input structure of an industry that remains unchanged regardless of its product mix, meaning that the technology assigned to the production of secondary products of an industry depends on the industry where it is produced (Miller & Blair, 2009; Raa & Rueda-Cantuche, 2007).

To start this approach each element of the use table (u_{ij}) is divided by the total output of industry j (g_i) and each element of the supply table (m_{ij}) is divided by the total demand of product i (q_i), to obtain the partitioned matrix D:

D $= \begin{bmatrix} 0 & Q \\ S & 0 \end{bmatrix}$, where each element of Q is given by $\frac{u_{ij}}{g_j}$ and each element of S is obtained by $\frac{m_{ij}}{q_j}$.

From D, considering the final demand aggregated into a single vector and then employing the general formulas for computing the inverse matrix it is possible to obtain the expression (1) (for further details see Miller & Blair, 2009).

$$\begin{bmatrix} I & -Q \\ -S & I \end{bmatrix}^{-1} = \begin{bmatrix} (I - QS)^{-1} & (I - QS)^{-1}Q \\ S(I - QS)^{-1} & I + S(I - QS)^{-1}Q \end{bmatrix} \tag{1}$$

From the rectangular IO model, it is possible to derive the expression (3) which is analogous to the Leontief inverse matrix, delivering an industry-by-commodity total requirements table, representing the total (direct and indirect) variation of each impact considered from industry j caused by the variation of one unit of final demand of commodity i (Miller & Blair, 2009; Locker et al. 2009). First, it is necessary to calculate the direct impact coefficients R, where each element,r_{kj}, is the amount of impact k produced per monetary unit of industry j's output (Hendrickson et al., 1998, 2006; Marques et al., 2006). Hence, the level of impacts associated with a given vector of total outputs is expressed in the expression (2) where **r** is the vector of impact levels:

$$\boldsymbol{r} = R\boldsymbol{x} \tag{2}$$

Consequently, when parameter x_i in Eq. (2) is replaced by the equation presented on the inferior left side of (1) it is obtained:

$$\boldsymbol{r} = R\left[S(I - QS)^{-1}\right]\boldsymbol{y} \tag{3}$$

In the assessment of the domestic impacts directly linked to each LC phase, the SUT format at basic prices was used, removing the imports.

4 Results

In this section, the main results found are presented in Tables 2 and 3 and discussed hereafter.

In this section, the main results found are presented in Tables 2 and 3 and discussed hereafter.

Starting with the performance of our reference building, its annual total primary energy consumption (TPEC) can go up to 0.15 TJ, the GHG emissions can achieve 5.24 tonnes of CO2eq, the costs related to energy supply and environmental impacts, exempted of taxes can reach 1937€, the household energy bill can attain 3722€, and the potential number of premature deaths caused by particulate matter emissions can achieve about 6.48E-09. Regarding GVA and employment impacts created during the MPIM phases, since none of the business as usual (BAU) technologies in place will be produced again to be part of the energy efficiency packages, their value is null.

According to our analysis, results show that the solutions which exclusively employ extruded polystyrene (XPS) for roof insulation present the highest SIR and the lowest GPBT and EPBT. Package nº 151, which only adds 40 mm of thickness'

Table 2 Impacts of EE packages accounting for biomass CO2 emissions

Package	Façades		Roof		Windows	SHC	DHW	Solar heater	GHG savings	GPBT	EPBT	SIR	NPV	Jobs	GVA	Impact on Household budget	Premature deaths	Energy Savings
Nº	Type	m	Type	m	Type	Type	Type	Yes /No	Tons CO2eq	Days	Days	Ratio	€	Nº	€	%	Nº	TJ
151	0	0	XPS	0.04	SAW	EH	GWH	No	1.81	0.335	0.190	23.83	654.4	0.000198	5.76	11.96	2.09E-09	0.0537
265	0	0	XPS	0.06	SAW	EH	GWH	No	1,98	0.327	0.185	20.68	710.5	0.000207	6.09	13.07	2.28E-09	0.0587
10,965	XPS	0.08	XPS	0.08	SAW	AHP	GWH	Yes	4.51	7.449	2.874	4.20	1301.0	0.00285	103	29.38	1.50E-09	0.1318
18,150	XPS	0.12	XPS	0.10	SAW	AHP	EWH	No	4.17	3.270	1.446	4.42	1196.6	0.0016	57.1	26.35	4.18E-09	0.1191
18,331	XPS	0.12	XPS	0.12	DPW	BB		Yes	2.85	21.589	5.234	2.38	913.7	0.00484	174	30.79	−2.64E-07	0.1290
18,335	XPS	0.12	XPS	0.12	DPW	DHW + H + C HP		Yes	4.62	13.717	5.758	1.75	757.9	0.00696	244	30.27	−4.05E-09	0.1350
18,339	XPS	0.12	XPS	0.12	DPW	AHP	DHP	Yes	4.64	9.825	3.876	2.59	1079.9	0.00424	151	30.06	−1.05E-09	0.1348
91,257	ICB	0.12	ICB	0.12	SAW	DHW + H + C HP		Yes	4.54	17.219	7.316	1.15	226.45	0.00997	402	30.03	−6.20E-09	0.1334

Source Authors' own elaboration

Table 3 Impacts of EE packages assuming carbon-neutrality of biomass

Package	Façades		Roof		Windows	SHC	DHW	Solar heater	GHG savings	GPBT	EPBT	SIR	NPV	Jobs	GVA	Impact on Household budget	Premature deaths	Energy Savings
Nº	Type	m	Type	m	Type	Type	Type	Yes /No	Tons CO2eq	Days	Days	Ratio	€	Nº	€	%	Nº	TJ
151	0	0	XPS	0.04	SAW	EH	GWH	No	1.81	0.335	0.190	23.83	654.4	0.000198	5.76	11.96	2.09E-09	0.0537
265	0	0	XPS	0.06	SAW	EH	GWH	No	1,98	0.327	0.185	20.68	710.5	0.000207	6.09	13.07	2.28E-09	0.0587
10,965	XPS	0.08	XPS	0.08	SAW	AHP	GWH	Yes	4.51	7.449	2.874	4.20	1301.0	0.00285	103	29.38	1.50E-09	0.1318
18,150	XPS	0.12	XPS	0.10	SAW	AHP	EWH	No	4.17	3.270	1.446	4.42	1196.6	0.0016	57.1	26.35	4.18E-09	0.1191
18,254	XPS	0.12	XPS	0.12	SAW	BB		No	5.08	7.008	3.133	3.82	1164.6	0.00275	99.8	29.98	−3.7E-07	0.1233
18,331	XPS	0.12	XPS	0.12	DPW	BB		Yes	5.02	12.572	5.234	2.52	1004.9	0.00484	174	30.79	−2.64E-07	0.1290
18,335	XPS	0.12	XPS	0.12	DPW	DHW + H + C HP		Yes	4.62	13.717	5.758	1.75	757.9	0.00696	244	30.27	−4.05E-09	0.1350
91,257	ICB	0.12	ICB	0.12	SAW	DHW + H + C HP		Yes	4.54	17.219	7.316	1.15	226.45	0.00997	402	30.03	−6.20E-09	0.1334

Source Authors' own elaboration

Note XPS—Extruded polystyrene; ICB—Insulation cork board; SAW—Single-glazed aluminium window; DAW—Double-glazed aluminium window; DPW—Double-glazed PVC window; EH—Electric heater; GWH - Gas water heater; EWH—Electric water heater; DHW + H + C HP—DHW + heating and cooling Heat pump; BB—Biomass boiler; AHP Air-source Heat Pump; DHP—DHW heat pump

XPS, allows the annual savings to exceed the annualized investment 23.83 times, while package 265, which only adds XPS with 60 mm of thickness, minimizes the time needed to recover the embodied energy and GHG emissions in the MPIM phases, up to 0.185 days and 0.327 days, respectively. Regarding packages that combine roof and façades insulation with space heating and cooling and DHW systems, package 10,965 composed of XPS with 80 mm of thickness for roof and façades insulation and a heat pump for replacing the electric heater for space heating is the solution with the highest annualized NPV that can go up to 1301.00€. Solution 18,150 allows reducing up to 65.23% of the potential number of premature deaths and consists in adding XPS with 120 mm and 100 mm of thickness to the façades and roof, respectively, by a heat pump to replace the electric heater and by an electric water heater to substitute the gas-fired water heater. The solution that maximizes the household budget is package 18,331 composed of XPS with 120 mm of thickness for insulation of roof and façades, double-glazed PVC frame windows to replace the single-glazed aluminum windows in use, and a biomass boiler for space heating and DHW. This solution allows for increasing the budget of a family in the poverty risk threshold by up to 31%. Package 18,335 made of XPS with 120 mm of thickness for insulation of roof and façades, double-glazed PVC frame windows, and a heat pump for space heating and cooling and DHW, is the solution that reaches the highest energy savings of about 0.1350 TJ. Package 91,257, which includes adding insulation cork board (ICB) with a thickness of 120 mm for roof and façade insulation and a heat pump for space heating and cooling as well as DHW, is the greatest way to increase economic and labor benefits. With this method, 0.00997 annualized full-time equivalent employment can be produced in the MPIM stages, and the GVA can increase to 402.00€. Finally, the package that maximizes the GHG savings (4.64 tons of CO2eq) is the 18,339 composed of XPS with 120 mm of thickness for insulation of roof and façades, double-glazed PVC frame windows, and a heat pump for space heating and cooling and a DHW heat pump for DHW.

After describing the best performances in each impact attained by the packages considered is important for the DMs to understand that trade-offs are always needed for selecting the best EE solutions to be funded. Therefore, taking into account the annualized SIR of the remaining packages this value varies from 1.15 to 20.68. In terms of TPES and GHG savings, the variation of these impacts can range from 0.0537 TJ to 0.1348 TJ, and from 1.81 tonnes of CO2eq to 4.62 tonnes of CO2eq, respectively. This results in an EPBT that can range from 0.190 to 7.32 days and a GPBT that can range from 0.335 to 17.21 days. The annualized NPV can change between 226.45€ and 1196.60€. During the MPIM phases, the employment and GVA creation values per year of the technologies' lifespan can range between 0.000198 and 0.00696, or 5.76€ and 244.00€, respectively. By using packages 151, 265, and 10,965, the potential number of premature deaths can be decreased by 32.61, 35.58, and 23.41%, whereas packages 18,331, 18,335, 91,257, 18,339, and 18,254 will result in an increase of up to 4119.57%, 63.20%, 96.75%, 16.38%, and 5773.64%, respectively. Finally, using packages 18,335, 91,257, 18,339, and 18,254, the household budget can increase by up to 30%.

If carbon-neutrality of biomass is considered the solution that allows achieving the highest GHG savings, changes from package 18,339 to package 18,254 composed of XPS with 120 mm of thickness for insulation of roof and façades, double-glazed PVC frame windows, and a biomass boiler for space heating and DHW. This solution can maximize the GHG savings by up to 5.08 tons of CO2eq. Additionally, in package 18,331 GHG emissions and the annualized GPBT decrease by 2.17 tons of CO2eq and about 9 days, respectively, while SIR can increase up to 0.14 and NPV reach 91.20€.

The findings show that the methodology outlined in this study should be supported in the decision-making process for the funding of EE measures because it allows for the development of a comprehensive evaluation of the impacts of investing in the technologies being examined, integrating the manufacturing, packaging, installation, and maintenance phases with the operation phase. Another benefit of this methodology is its ability to interact with different methodologies for assessing the energy needs of buildings, such as the seasonal approach employed in this study or the dynamic simulation method. In addition to energy savings and GHG emissions, the use of IO methodologies also enables the examination of several other impacts and benefits that are crucial for decision-making when designing new programs to finance EE.

5 Conclusions and Further Research

This paper presents a novel methodological approach that integrates an HIO-LCA framework with the Portuguese seasonal method employed by the Portuguese building energy certification system for calculating the energy performance of buildings, to evaluate the energy, economic, environmental, and social benefits/impacts of investing in distinct EE packages in the Portuguese residential sector. These packages have been created through the combination of thermal insulation options for roofs and façades, with the replacement of windows, also considering the use of space heating and cooling and DHW systems using a solar collector as a backup. Through this methodology different lifecycle phases are considered, in addition to operation, the SUT structure is employed instead of the symmetric format for reasons of more comprehensiveness, and DMs are supported to design suitable EE funding policies. This methodology was tested using the characteristics of a T2 single dwelling built between 1961 and 1991, located in Coimbra.

Our findings suggest that package 151 presents the highest SIR; package 265 minimizes the EPBT and GPBT; package 91,257 is the best solution to increase the economic and labor benefits; package 18,339 is the "cleanest" one; package 10,965 has the highest annualized NPV; package 18,331 maximizes the household budget of a family under poverty risk; package 18,150 has the highest potential to reduce premature deaths, and solution 18,335 reaches the highest energy savings. If carbon-neutrality is assumed for biomass, the minimization of GHG savings is obtained with solution 18,254.

Given these trade-offs, it is evident that multiobjective optimization methods that incorporate the impacts and benefits assessed by the methodology described in this study are essential for assisting DMs in modeling different investment strategies and designing future EE programs that reflect their preferences. Further research is also expected to cover the assessment of other Portuguese locations, other sorts of impacts (i.e., on public budget and energy poverty), and the consideration of the end-of-life phase.

Acknowledgements This research was supported by the doctoral Grant SFRH/BD/151353/2021 financed by the Portuguese Foundation for Science and Technology (FCT), under MIT Portugal Program, and by the project grants UIDB/00308/2020 and T4ENERTEC (POCI-01-0145-FEDER-029820) co-funded by ERDF - European Regional Development Fund through Operational Program for Competitiveness and Internationalization - COMPETE 2020 and by the Portuguese Foundation for Science and Technology.

References

Asadi, E., Da Silva, M. G., Antunes, C. H., & Dias, L. (2012). Multi-objective optimization for building retrofit strategies: A model and an application. *Energy and Buildings, 44*, 81–87.

Bilec, M., Ries, R., Matthews, H. S., & Sharrard, A. L. (2006). Example of a hybrid life-cycle assessment of construction processes. *Journal of Infrastructure Systems, 12*(4), 207–215. https://doi.org/10.1061/(ASCE)1076-0342(2006)12:4(207)

Cellura, M., Di Gangi, A., Longo, S., & Orioli, A. (2013). An Italian input-output model for the assessment of energy and environmental benefits arising from retrofit actions of buildings. *Energy and Buildings, 62*, 97–106. https://doi.org/10.1016/j.enbuild.2013.02.056

Crawford, R. H. (2009). Life cycle energy and greenhouse emissions analysis of wind turbines and the effect of size on energy yield. In: Renewable and Sustainable Energy Reviews (Vol. 13, Issue 9, pp. 2653–2660). https://doi.org/10.1016/j.rser.2009.07.008

de Carvalho, A. L., Antunes, C. H., Freire, F., & Henriques, C. O. (2016). A multi-objective interactive approach to assess economic-energy-environment trade-offs in Brazil. *Renewable and Sustainable Energy Reviews, 54*, 1429–1442 S1364032115011430. https://doi.org/10.1016/j.rser.2015.10.064

Energy Observatory, DGEG, & ADENE. (2021). Energy in numbers–edition 2021. ADENE–Agência para a Energia.

European Commission. (2019). Clean energy for all Europeans package|Energy. In European Commission. https://ec.europa.eu/energy/topics/energy-strategy/clean-energy-all-europeans_en

European Commission. (2020). A renovation wave for Europe–greening our buildings, creating jobs, improving lives, *vol. COM (2020)*, (Issue 662). https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52020DC0662.

European Commission. (2021). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "Fit for 55": delivering the EU's 2030 Climate Target on the way to climate neutrality. COM/2021/550 Final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021DC0550

Hendrickson, C., Horvath, A., Joshi, S., & Lave, L. (1998). Economic input-output models for environmental life-cycle assessment. *Environmental Science and Technology, 32*(7), 184A-191A. https://doi.org/10.1021/es983471i

Hendrickson, C. T., Horvath, A., Joshi, S., Klausner, M., Lave, L. B., & McMichael, F. C. (1997). Comparing two life cycle assessment approaches: a process model- vs. economic input-output-based assessment. In: IEEE International Symposium on Electronics & the Environment, 176–181. https://doi.org/10.1109/isee.1997.605313

Hendrickson, C. T., Lave, L. B., & Matthews, H. S. (2006). Environmental life cycle assessment of goods and services: An input-output approach. In Environmental Life Cycle Assessment of Goods and Services: An Input-Output Approach (1st Editio). Routledge. https://doi.org/10.4324/9781936331383.

Henriques, Carla Oiveira, & Coelho, D. H. (2017). A multiobjective interval portfolio formulation approach for supporting the selection of energy efficient lighting technologies. In: International Proceedings of Chemical, Biological and Environmental Engineering (ICEEB2017). https://doi.org/10.7763/IPCBEE

Henriques, C. O., Coelho, D. H., & Neves, M. E. D. (2020). Investment planning in energy efficiency programs: A portfolio based approach. *Operational Research*, 1–35,. https://doi.org/10.1007/s12351-020-00566-6

Horowitz, K. J., & Planting, M. A. (2006). Concepts and Methods of the U.S. Input-Output Accounts. In Bureau of economic analysis. U.S. Department of Commerce.

INE. (2012). Censos 2011: XV Recenseamento Geral da População e V Recenseamento Geral da Habitação. INE Statistics Portugal.

INE.(2017). Sectoral Waste. https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0006061&contexto=bd&selTab=tab2

INE. (2019). Estatísticas da Construção e Habitação - 2018. In INE Statistics Portugal. www.ine.pt

Locke, G., Blank, R. M., Landefeld, J. S., & Marcuss, R. D. (2009). Measuring the Nation's economy: an Industry Perspective.

Marques, J. L, Martins, J. M., & Anselmo Castro, E. (2006). Análise input-output rectangular inter-regional: emissões de CO2 em Portugal e o protocolo de Quioto. Revista Portuguesa de Estudos Regionais, 11, 25–59. https://www.redalyc.org/pdf/5143/514352348002.pdf

Miller, R. E., & Blair, P. D. (2009). Input–output analysis: Foundations and extensions, 2nd edn. Cambridge University Press. www.cambridge.org/9780521517133%0D

Ministry of Economy and Employment. (2013). Decree Law No. 118/2013, of 2013–08–20 - Energy Performance Regulation of Residential Buildings. In Diário da República n.o 159/2013, Série I de 2013–08–20 (pp. 4988–5005). https://dre.pt/pesquisa/-/search/499237/details/maximized

Ministry of Planning. (2021). Recovery and resilience plan–recovering Portugal, building the future. https://www.portugal.gov.pt/download-ficheiros/ficheiro.aspx?v=%3D%3DBQAAAB%2BLCAAAAAAABAAzNDQzNgYA62SpeQUAAAA%3D.

OECD. (2017). National and Social Accounts. https://stats.oecd.org/

Oliveira, C., Coelho, D., & da Silva, P. P. (2014). A prospective analysis of the employment impacts of energy efficiency retrofit investment in the Portuguese building stock by 2020. *International Journal of Sustainable Energy Planning and Management, 2*, 81–92. https://doi.org/10.5278/ijsepm.2014.2.7

Pinto, A., & Fragoso, R. (2018). Indicador de comportamento térmico solar passivo de habitações—Estação de aquecimento. www.lnec.pt

Presidency of the Council of Ministers. (2013). National Energy Efficiency Action Plan (NEEAP 2016). In Resolução do Conselho de Ministros n.o 20/2013, Diário da República, 1a Série–no 70 de 10 abril de 2013.

Presidency of the Council of Ministers. (2020). Resolution of the Council of Ministersno. 53/2020 - National Energy and Climate Plan 2030 (NECP 2030). Diário da República n.o 133/2020, Série I de 2020–07–10, 2–158.

Presidency of the Council of Ministers. (2021). Resolution of the Council of Ministers No. 8-A/2021 - Long Term Strategy for the Renovation of Buildings. Diário Da República n.o 23/2021, 1o Suplemento, Série I de 2021–02–03, 16-(2) to 16-(105). https://data.dre.pt/eli/resolconsmin/8-A/2021/02/03/p/dre

Raa, T. T., & Rueda-Cantuche, J. M. (2007). Stochastic analysis of input-output multipliers on the basis of use and make tables. *Review of Income and Wealth*, *53*(2), 318–334. https://doi.org/10.1111/j.1475-4991.2007.00227.x

Reuter, M., Patel, M. K., Eichhammer, W., Lapillonne, B., & Pollier, K. (2020). A comprehensive indicator set for measuring multiple benefits of energy efficiency. *Energy policy, 139*, 111284. https://doi.org/10.1016/j.enpol.2020.111284

Ryan, L., & Campbell, N. (2012). Spreading the net: the multiple benefits of energy efficiency improvements.

Sargento, A., Ramos, P. N., & Hewings, G. (2011). Input-output modelling based on total-use rectangular tables: is this a better way?. *Notas Económicas, 34*. https://doi.org/10.14195/2183-203X_34_1

Säynäjoki, A., Heinonen, J., Junnonen, J. M., & Junnila, S. (2017). Input–output and process LCAs in the building sector: Are the results compatible with each other? *Carbon Management, 8*(2), 155–166. https://doi.org/10.1080/17583004.2017.1309200

Singh, V. K., Henriques, C. O., & Martins, A. G. (2018a). Fostering investment on energy efficient appliances in India–A multi-perspective economic input-output lifecycle assessment. *Energy, 149*, 1022–1035. https://doi.org/10.1016/j.energy.2018.01.140

Singh, V. K., Henriques, C. O., & Martins, A. G. (2018b). A Multiperspective Assessment of Best Available Energy End-Use Technologies in India's Households. *Process Integration and Optimization for Sustainability, 3*(1), 89–99. https://doi.org/10.1007/s41660-018-0054-1

Singh, V. K., Henriques, C. O., & Martins, A. G. (2019). A multiobjective optimization approach to support end-use energy efficiency policy design—The case-study of India. *International Journal of Sustainable Energy Planning and Management, 23*, 55–68. https://doi.org/10.5278/ijsepm.2408

Stephan, A., & Stephan, L. (2014). Reducing the total life cycle energy demand of recent residential buildings in Lebanon. *Energy*, *74*(C), 618–637. https://doi.org/10.1016/j.energy.2014.07.028.

Strømman, A. H., Peters, G. P., & Hertwich, E. G. (2009). Approaches to correct for double counting in tiered hybrid life cycle inventories. *Journal of Cleaner Production, 17*(2), 248–254. https://doi.org/10.1016/j.jclepro.2008.05.003

Suh, S. (2006). Reply: Downstream cut-offs in integrated hybrid life-cycle assessment. *Ecological Economics*, *59*(1), 7–12. Elsevier. https://doi.org/10.1016/j.ecolecon.2005.07.036.

Tadeu, S., Tadeu, A., Simões, N., Gonçalves, M., & Prado, R. (2018). A sensitivity analysis of a cost optimality study on the energy retrofit of a single-family reference building in Portugal. *Energy Efficiency, 11*(6), 1411–1432.

Tenente, M., Henriques, C., Gomes, Á., da Silva, P. P., & Amado, L. (2021). A novel approach for assessing the multiple benefits of energy efficiency technologies. In: European Council for an Energy-Efficient Economy—Eceee 2021 Summer Study on Energy Efficiency, Eceee Summer Study Proceedings, pp. 1065–1075.

Woolf, T., Neme, C., Kushler, M., Schiller, S. R., Eckman, T., & Michals, J. (2017). National Standard Practice Manual for Assessing Cost-Effectiveness of Energy Efficiency Resources. https://nationalefficiencyscreening.org

Zhan, J., Liu, W., Wu, F., Li, Z., & Wang, C. (2018). Life cycle energy consumption and greenhouse gas emissions of urban residential buildings in Guangzhou city. *Journal of Cleaner Production, 194*, 318–326. https://doi.org/10.1016/j.jclepro.2018.05.124

Case Studies

EU Operational Programmes Reporting: From Basics to Practices

Ana Amaro, Carla Henriques, and Clara Viseu

Abstract We examine the main existing challenges that currently arise in the assessment of European Union (EU) funds devoted to three thematic objectives (TOs): Research and Innovation (R&I); Low-carbon economy (LCE); and Information and Communication Technologies (ICT). In this regard, a literature review on the European Regional Development Fund (ERDF) initiatives is performed, with a special focus on the Portuguese (PT) case, also addressing their assessment and reporting practices. Data systematization is coupled with the European Commission (EC)'s main guidelines and with the guidance recommendations brought by Management authorities (MA) for the 2014–2020 period. A bibliometric analysis is conducted to further understand the current research interest in the evaluation of EU funds, and the type of assessment methods and reporting practices employed. Most of the approaches rely on cost–benefit analysis and place less attention on data availability, variable selection, and monitoring/assessment options. The selection and application of the framework indicators, either related to their financial execution or achievement, are assumed as critical factors concerning the monitoring, reporting, and assessment processes. Our findings emphasize the need for harmonization and simplification of the reporting techniques, also highlighting the sparse data availability and some reporting conflicts.

A. Amaro (✉)
CEGIST-IST-UL, 1049-001 Lisboa, Portugal
e-mail: aamaro@iscac.pt

A. Amaro · C. Henriques · C. Viseu
Polytechnic of Coimbra, Coimbra Business School, ISCAC, 3045-601 Coimbra, Portugal

C. Henriques
DEEC, INESC Coimbra, Rua Sílvio Lima, Polo II, 3030-290 Coimbra, Portugal

Centre for Business and Economics Research (CeBER), Faculty of Economics, University of Coimbra, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

C. Viseu
CMA-UBI, 6200-001 Covilhã, Portugal

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_10

Keywords ERDF · Portuguese operational programs · Reporting · Indicators · Performance

1 Introduction

Despite the prolific number of studies investigating the effects of EU funds, empirical evidence shows mixed, and, sometimes, contradictory, results (Gramillano et al., 2018a, b). While some authors found a positive effect of EU funds on the economic growth of regions (Puigcerver-Penalver et al., 2007), others highlighted the existence of a maximum desirable level of funds transfer, beyond which the funds may increase regional gaps within countries (Kyriacou & Roca-Sagales, 2012). Also, some contributions acknowledged no statistically significant impact of Cohesion Funds (CF) on convergence, underlining how disparities persist in the EU (Dall'Erba & Le Gallo, 2008).

As Scotti et al. (2022) emphasized, although the EU Cohesion Policy has progressively diversified the sectors targeted for funding, with possible heterogeneous impacts on local growth, the literature is still largely oriented to the analysis of aggregate impacts. Thus, these authors proposed a granular investigation of the sectoral impacts of Structural and CF on European NUTS 2 regions for the period 2007–2014 and concluded that expenditures in energy, Research and Development (R&D), and transportation sectors stimulated a higher Gross Domestic Product (GDP) per capita, consistently reducing production costs, increasing accessibility and innovation in the beneficiary regions.

Moreover, the assessment of funded programs is still a largely discussed topic. In this regard, the EC identified two main assessment problems with the system of indicators (Nigohosyan & Vutsova, 2018): difficulties in establishing the cause-and-effect relationships between actions, results, and impacts, due to the influence of external factors; and difficulties in measuring the impacts, because they are usually the cumulative effect of many actions, affect diverse populations, and it takes time for them to show their actual effects. For the 2014–2020 period the EC tried to overcome these difficulties in the ERDF context by discarding the concept of impact indicators and introducing a new intervention logic. Nevertheless, Nigohosyan & Vutsova (2018) claimed that despite the changes, some of the identified problems with the indicators were not solved. Besides, these authors suggested that ongoing evaluations should include activity's in-depth analysis of the relevance of the selected output indicators. All in all, Nigohosyan & Vutsova (2018) consider that the introduced changes are well-intended, but they still need to be perfected. The system of common output indicators should be re-examined, while the requirements for the result indicators should not be contradictory. The current practices show that even solid and logical principles, such as the fact that EU policies should reflect economic, social, political, and institutional differences to maximize both the local and the aggregate potential for economic growth (Barca & McCann, 2012), might not be sufficient when the different countries select indicators.

Furthermore, in the current programming period (2021–2027), the lessons learned from other programming experiences within and beyond the European Structural and Investment Fund (ESIF) framework can make a difference, since ESIF operates in very different local contexts and handles very heterogeneous economic and social regional environments. In this framework, a previous study by Bachtler & Wren (2006) reported that even if Cohesion Policy had a unified regulatory framework, it should address different national and regional circumstances embedded in a variety of institutional arrangements, bearing in mind that its operations comprise a multiplicity of measures and a diversity of national, regional, and local rules and systems.

Additionally, programs consist of a range of interventions (physical and economic infrastructures, business and technological developments, human resources, innovation, and environmental improvement) based on a mix of financial instruments for many types of beneficiaries. This variety of targets and contextual conditions is per se a challenge for any evaluation exercise (Henriques & Viseu, 2022a, b).

In this vein, Nigohosyan & Vutsova (2018) emphasized that the EC should try, in the next programming period (i.e., 2021–2027), to unify as much as possible the understanding of indicators under the different EU funds. An easy step would be the inclusion of all indicators into a single guidance document. Their study assessed the possibility of expanding the current list of common output indicators and the feasibility of developing a list of common direct result indicators for post-2020 ERDF and CF interventions.

Regarding the allocation of the EU budget, Gramillano et al. (2018a, b) suggested that spatial and sectoral effects can contribute to the design of more effective distribution of EU funds. Also, due to the wide variety of projects financed by CF, the EU Regional Cohesion Policy has been defined as a "do it all policy". Hence, policymakers are currently focused on the economic impact of investments across different sectors, since heterogeneous levels of local development may be achieved depending on the economic activity in which the EU transfers are allocated (Cortuk & Guler, 2015). Indeed, investments in certain sectors might have immediate positive effects, while other types of investments might generate a significant impact only in a long-term perspective (Scandizzo et al., 2020).

Moreover, the magnitude of economic multipliers might be different across sectors and dependent on the level of diversification and complementarity of expenditures (Auerbach & Gorodnichenko, 2012; Duranton & Venables, 2021).

With the foregoing in mind, we investigate the current major issues inherent to the assessment of ESIF committed to three TOs: R&I, LCE, and ICT. A literature review on ERDF implementations is conducted, with a specific focus on the PT situation, as well as its evaluation and reporting practices.

The research questions that we want to address are the following:

RQ1: "What are the main areas of concern of the studies devoted to the assessment of ERDF?"

RQ2. "What are the main challenges inherent to the selection of the variables/indicators in the assessment of ERDF?"

RQ3. "What are the best practices found in the assessment of ERDF?"

RQ4. "What are the main gaps found regarding the methodologies used to assess ERDF?"

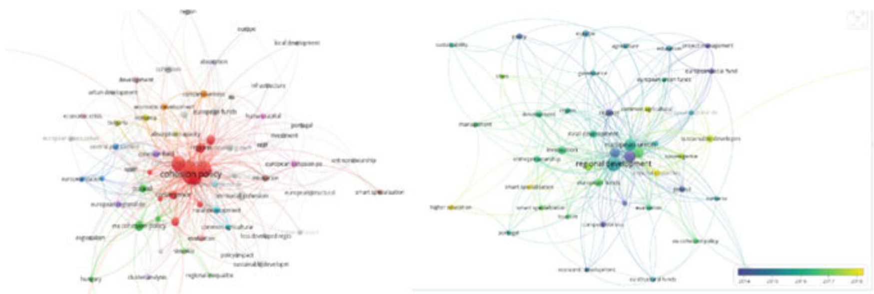
This paper is structured as follows. In Sect. 2 a systematized literature review is conducted. Section 3 goes through the Portuguese case. Section 4 describes the main results found for the Portuguese assessment and reporting systems. Finally, some conclusions are presented, and future work developments are indicated.

2 Literature Review

A systematic literature review has been done on the *Web of Knowledge* database, using a set of keywords combined with Boolean operators, and it covers a wide number of data sources (i.e., scientific journals, books, proceeding papers, etc.). Data collected from EU Reporting were also attained in the analysis. The bibliographic results were saved in data files and comprise the search over all the publication content. These text files were collected and manipulated using Vosviewer (i.e., https://www.vosviewer.com/) to map the bibliographic content—see Figs. 1 and 2.

By using the keywords "EU funds", the search returned 28,809 references that were filtered regarding the type of publication and then coupled with our focal interest, that is the "ERDF". The latter generated 1,200 references, and while combined conjunctively (i.e., using the "AND" operator) with the "EU funds" keyword returned 532 references. These were extracted from the database to be mapped by Vosviewer. Besides, another collection of references was obtained by adding "Portugal" and "Funds" to the "ERDF" keyword.

As can be observed in Fig. 1a the research published so far regarding the EU funds mostly focus on the cohesion policy and, entering a deep detail, "structural funds" and "regional development" appear as commonly referenced keyword, as, mapped in Fig. 2a.



(a) (b)

Fig. 1 Bibliometric clustering **a** and timeline **b** for EU combined with ERDF and funds. *Source* Authors' own elaboration

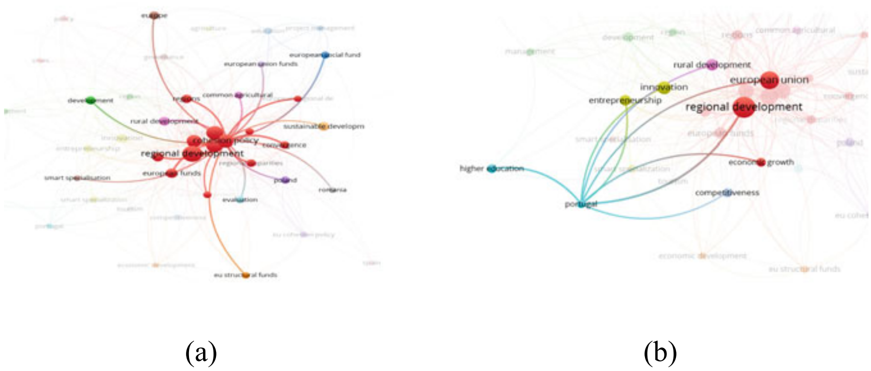


Fig. 2 Bibliometric clustering with highlight focus on **a** central thematic and **b** PT-related arcs. *Source* Authors' own elaboration

While introducing the timeline into the data collection (Fig. 1b), the global cluster results show that most references are up to 2016 (green). Also, more recent publications depict a higher concern over the following themes: "sustainability", "smart specialization", "high education", or "regional disparities" (yellow). These are topics quite prominent in today's EU political agenda and are in line with the latest worldwide concerns.

Finally, regarding Portugal (Fig. 2b), most of the studies address "regional development", but "competitiveness", "higher education", "innovation" or "rural development" are also part of the connecting arcs.

Subsequently, the literature review was systematized by grouping the references into four categories: Data and variables selection; Indicators and monitoring; Best Practices; and Methods.

2.1 Data and Variables Selection

The lack of data and the heterogeneous definitions of relevant indicators further complicate the analysis of the Operational Programs (OPs) funded by ERDF (Henriques et al., 2022a, b). Both, policy and economic performance/outcome indicators can be measured/proxied by different variables and the choice of these proxies may have important implications for the results of the various analyses (Pastor et al., 2010). In most cases, the policy variables under study are 'payments' or 'commitments' and 'GDP Growth Rate Per Capita' or 'Employment Rate', which are used as proxies for economic performance. Therefore, the choice of the policy variables can be a determinant factor for the design and results of empirical analyses. For example, the use of actual policy 'expenditure' data instead of 'commitments', means having to consider the duration of the entire OP (Crescenzi & Giua, 2016).

Additionally, the data available to perform the assessments varies considerably with the type of TO under scrutiny. In the case of R&I (TO1), policymakers face additional challenges in the assessment of R&I policies, notably because of the scarcity of suitable data (Henriques et al., 2022a; Ortiz & Fernandez, 2022). The sparsity of regional-level research and data on ICT (TO2) at the firm level has also been highlighted by Ruiz-Rodríguez et al. (2018), Reggi and Gil-Garcia (2021) and Henriques and Viseu (2022a, b). For example, despite the ample set of data available in the latest European Innovation Scoreboard (Hollanders, 2021) with some indicators on ICT (e.g., Digital skills and business sector ICT specialists), Henriques & Viseu (2022a, b) only managed to consider three indicators for assessing the OPs related to boosting ICT adoption in SMEs (i.e., operations supported, eligible costs decided, and eligible spending). An additional difficulty refers to the identification of EU ICT targeted investments (Sörvik & Kleibrink, 2016). Aside from being an economic sector, ICT is still an essential portion of many other activity-related sectors (for example, e-Health) and a tool to assist other activities. Since ESIF actions might have multiple aims, it can be hard to pinpoint the ICT-related activities within the designated categories when the OPs are planned. The OPs' financial metrics are organized into intervention categories, TOs, and priority domains. Moreover, although the EU guidelines advocate that planned ICT initiatives should be classified primarily underneath TO2, these obtain funds under distinct TOs, and they are also integrated into various smart specialization strategies. For example, to consider the ICT SMEs support, there are only two dimensions of intervention that can be considered (Henriques & Viseu, 2022a, b; Reggi & Gil-Garcia, 2021; Sörvik & Kleibrink, 2016) under codes 4 and 82, that correspond to €1.7 billion and €304 million of planned investments, respectively (Sörvik & Kleibrink, 2016). These totals are available under multi-TO (€810 million), TO2 (€790 million), and TO3 (€349 million) and to a smaller level under TO1 and TO8 (Sörvik & Kleibrink, 2016). As a direct consequence, national and regional policymakers should use additional specific criteria that account for ICT results; tag expenditure that falls under other TO (rather than TO2) but has an ICT component; enhance the quality and completeness of ICT performance data at the regional and SME levels; and unify different data from diverse data sources (Henriques & Viseu, 2022a, b).

Furthermore, as noticed by Henriques et al., (2022a, b), despite the performance framework providing a set of implementation indicators, the data provided is frequently incomplete, and as a result, assessments end up considering a limited number of indicators and OPs. Finally, it is not possible to reach a complete match between the data obtained for the OPs' achievements and their financial implementation. This is especially true for the investment priority dedicated to SMEs (investment priority 4b), under TO4 (i.e., LCE), which is meant to increase energy efficiency and renewable energies and includes statistics for the accomplishment metrics but not for their financial implementation.

2.2 Indicators and Monitoring

Performance appraisal techniques, as well as the selection of indicators, provide critical insights into supported programs. The different types of indicators are based on the following aspects (European Commission, 2013): relationship with the variables (complete, partial, and complex indicators); information processing (elementary, derived, and compound indicators); information comparability (specific and common indicators); and information scope (context and program indicators). The primary categories of indicators (input, output, and result) are provided in all ESIF monitoring systems; however, there are significant differences. Apart from the division of indicators into quantitative and qualitative, the guidance documents for the ERDF and the European Social Fund (ESF) agree that indicators should be linked to the specific objectives and kept as close as possible to the activities. Regarding the result indicators, the major difference is that ERDF should not be limited to the supported entities since it addresses different local contexts and very heterogeneous economic and social regional realities, whereas ESF result indicators "capture the expected effects on participants or entities brought about by an operation" (European Commission, 2015). Therefore, the ERDF result indicators should capture the change in the Member States (MS), regions, areas, or affected populations because of a particular program. This leads to two main problems: difficulties in establishing a program's contribution to results and problems determining the target values.

Nonetheless, the results of Nigohosyan & Vutsova (2018) show that not all of the common indicators suggested by the EC follow its guidelines and the new intervention logic model could still lead to unclear logical links between activities, outputs, and results of the programs. These findings led to the recommendation for a mid-term review of the common ERDF indicator system.

Nigohosyan & Vutsova (2018) focused on the evolution of the ERDF indicators in the 2014–2020 programming period. The main question addressed was whether the new indicator system solved the problems of the past. These authors argued that despite the changes, the evolution of the ERDF indicators is incomplete, and it will likely be the reason for serious monitoring deficiencies and evaluation challenges.

A new understanding of 'impact' and the exclusion of impact indicators was put forward in the 2014–2020 guidance for evaluating the EU Cohesion Policy and the new ERDF-supported OPs (Nigohosyan & Vutsova, 2018).

Nigohosyan & Vutsova (2018) argued that despite the good justification for these changes, the 2014–2020 ERDF intervention logic and indicator system did not solve some of the well-known problems and brought new challenges to the evaluation of the ERDF-supported programs. The main challenges that remain to be solved are differing indicator concepts across the EU funds; inconsistency of the common output indicators; difficulties in establishing a program's contribution to results; persistent problems in determining the target values for results; and broad result indicators with an indirect link to interventions.

The same idea was reported by the EC (European Commission, 2018) clarifying the core differences between the period 2014–2020 and the ongoing Period 2021–2027. Supported on the key concept of simplification (i.e., more comparable data based on the use of fewer indicators), the 11 TO originated 5 Policy Objectives; the 3,573 Specific Objectives gave place to 21 Specific Objectives. Regarding the Results, the 5,082 program-specific indicators were replaced by 85 common indicators plus program specific; and on the outputs reporting side, 46 common indicators (6,481 records) and 4,813 program-specific indicators, were replaced by 85 common indicators plus program specific.

Also, in the 2014–2020 period, roughly half of the common ERDF output indicators could be viewed as another type of indicator (inputs or results). Indeed, some indicators can be viewed as output indicators in one context, and result indicators in another (e.g., the indicator "Reduction of greenhouse gases").

All along the developed research, Gramillano et al. (2018a, b), identified and emphasized the need for the harmonization and simplification of the used indicators. More comparable data based on the use of fewer indicators than those defined for the period 2014–2020 is desirable. Apart from the provided framework of common indicators, a match between the data gathered for the achievement indicators and the data from financial implementation is not fully feasible. Moreover, the data reported is often lacking (Henriques et al., 2022a, b). This is particularly true regarding the PT case, where the level of data details and the reporting practices frequently overlook common standards that hamper the programs evaluations and national OPs' comparability.

2.3 Best Practices

Many authors agree on the importance of "learning from the past". To make it feasible, some additional care must be placed on the chosen variables, the selected and monitoring indicators, and the implemented methodologies, since these are relevant to enhance that learning process.

Furthermore, to hasten the execution of the OPs, best practices from other MS should be examined, and administrative hurdles to obtaining funds should be reduced. MA should look for methods to boost project delivery by promoting the streamlining of payment request protocols and giving more guidance and support (Henriques et al., 2022b).

Frequently, assessment reports indicate that certain firms have revoked their subsidies, most likely owing to bank credit difficulties. In this respect, MA should be prepared to support businesses in securing additional financing options while also simplifying the conditions for attracting other institutional investors (Henriques et al., 2022b).

Gramillano et al. (2018a, b), presented a system of common indicators for the ERDF and CF interventions after 2020. The analysis performed covers the 11 TOs and is structured in two parts (I, for TOs 1, 3, 4, 5 and 6, and II for TOs 2, 7, 8, 9, 10 and

11). In this study, indicators (i.e., common output and direct result indicators) were evaluated according to their quality assessment, supported by key RACER principles (i.e., R (relevant), A (accepted), C (credible), E (easy to monitor), R (robust)) and other revised criteria (e.g., CREAM matrix assessment that sets out five criteria: Clear, Relevant, Economic and Available at a reasonable cost, adequate to provide information useful to assess performance and able to Monitor).

Moreover, time-bound indicators are also critical since they provide dates for measurement over time and monitoring is based on annual reporting or at least takes place at the end of the project. Besides, a debatable criterion should also be emphasized. Here, MA consultation collects information on the use of common output indicators to conduct benchmarking analyses or at least to use them in the future. This is to verify how much comparable information from 2014–2020 common outputs has been exploited since comparability is a major advantage of common indicators.

Globally, apart from the continuity of the best achievements from 2014–2020, the EC focused on: (i) the match of common outputs and results for interventions; (ii) simplification, harmonization, and data comparability; (iii) broader policy coverage; (iv) flexibility; (v) alignment with ESF; (vi) aggregation from project level; (vii) RACER—Financial Regulation criteria (Gramillano et al., 2018a).

Some clarification on the common ERDF and CF indicators was prompted by the EU (2018, Annex I and II), namely for the indicators to be selected in the programs, the data to be collected from projects via the monitoring systems, and the aggregated data reported by MS to the EC.

So, as suggested by Nigohosyan & Vutsova (2018), the EC should try in the next programming period to unify as much as possible the understanding of indicators under the different EU funds. An easy step would be the inclusion of all indicators into a single guidance document.

2.4 Methods

The literature evaluation identified desk research, monitoring data/data analysis, interviews, focus groups/facilitated workshops, surveys, and case studies as the major applied techniques to analyze ERDF TOs. Notwithstanding, the MS efforts to improve cohesion policy appraisal, only very few evaluations use more reliable methodologies, such as statistical methods or other mathematical techniques (Henriques et al, 2022a, b). Non-parametric approaches, like DEA, have turned into a noteworthy methodological alternative to the traditional approaches employed in similar contexts. The key benefit of utilizing this mathematical approach is the type of information that it can provide to MA on the inefficiency of the OPs when compared to their counterparts.

The benchmarks of inefficient OPs are also determined by DEA, and significant information about the best practices to follow to reach efficiency may be obtained. Nonparametric approaches, such as DEA, may readily manage many assessment criteria. Furthermore, DEA can help identify the key reasons that hamper efficiency, supplying policymakers with relevant knowledge on how to solve them. For example, Gómez-García et al. (2012) evaluated the efficiency of the implementation of ESIF allocated to TO1 in this context. Furthermore, Gouveia et al. (2021) used the Value-Based DEA method to evaluate the implementation of an ESIF aimed at enhancing the competitiveness of SMEs throughout multiple OPs (national and regional). In addition, Henriques et al. (2022b) evaluated 102 OPs from 22 EU nations dedicated to supporting an LCE in SMEs using the output-oriented variant of the slack-based measure (SBM) paired with cluster analysis. Lastly, Henriques et al. (2022a) evaluated 53 OPs from 19 MS committed to boosting R&I in SMEs using the non-oriented form of the network SBM approach in conjunction with cluster analysis.

Furthermore, DEA models are easily adaptable to evaluate different TOs if the basic rule proposed by Golany et al. (1989) is followed, namely, the number of DMUs (in this case, the OPs, EU funds, regions, countries, etc.) under evaluation should at least double the number of inputs and outputs (the indicators used in the evaluation). Though DEA offers undeniable benefits over other conventional methods (for example, microeconomic analyses that utilize control groups and case study analysis), there is currently a dearth of academic interest in its application in the context of ESIF efficiency appraisal.

This form of analysis is especially important if the programs are still in progress since it allows MA to predict the influence that prospective changes in output/input levels would have on the levels of efficiency attained by the OPs.

Unlike other approaches and methodologies that are specially applied for ex-post or ex-ante evaluation of cohesion policies, the DEA approach allows us to assess the efficiency of OPs' deployment across the programmatic time horizon (thus allowing us to perform midterm/terminal assessments), so that the required initiatives can be implemented within the time necessary for producing the appropriate changes during the timeframe in headway.

Due to the lack of more robust approaches during midterm/terminal assessments, the adoption of nonparametric methodologies can be particularly beneficial and suitable, mostly because the existing metrics for appraising the Cohesion Policy can be employed with other methodologies and contextual indicators. This can be done by combining this sort of analysis with Stochastic Frontier Approach (SFA), for example, thus allowing us to understand if the inefficient results obtained are mainly related to managerial failures or to the contextual environment or statistical noise (Henriques & Viseu, 2022a, b).

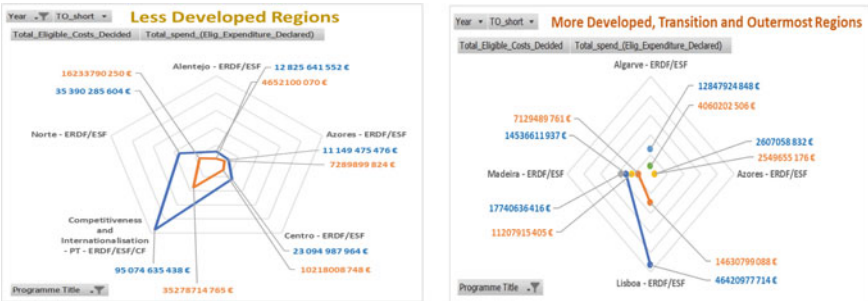
3 The Portuguese Case: Main Findings

For the PT OPs, different funding dimensions on ERDF were exploited and characterized. Apart from data collection, data curation (i.e., identification of missing and null values, data consistency, etc.), data characterization (i.e., evaluation of the main statistics), and some data visualizations were used to analyze the PT public dataset.

The ESIF comprises the allocation of €461 billion and is distributed by six funds—See Fig. 3a. In 2014, a partnership agreement (PA) between Portugal and the European Commission, called Portugal 2020 (P2020), was signed. Based on that partnership an ESIF budget of €25,860 million was assigned to Portugal, which represents about 5% of the ESIF for all MS and about 2,561€ per capita. The P2020 programming and implementation were divided into four thematic domains (Vaquero et al., 2020): competitiveness and internationalization, social inclusion and employment, human capital and sustainability and efficiency in the use of resources, and seven regional operational programs (OPs). The distribution of the ESIF funds shows that approximately 79% of the total amount assigned to Portugal, goes to the three funds included in the Cohesion Policies, that is, ERDF (45.3%), ESF (24.7%), and CF (9%) (European Commission, 2022).

In terms of eligibility for the ESIF (ERDF, CF, ESF, EAFRD, and EMFF), the seven Portuguese regions are divided into Less developed regions (GDP per capita <75% EU average): Norte, Centro, Alentejo, and the Azores, with a co-financing rate of 85%; Regions of transition (GDP per capita between 75 and 90%): Algarve, with a co-financing rate of 80%; and More developed regions (GDP per capita >90%): Lisboa and Madeira, with a co-financing rate of 50% to Lisboa and 85% to Madeira. Figure 3 depicts the execution of the EU funds in Portugal per type of region. As formerly stated, the programming and implementation defined by the PA for Portugal are based on four key thematic domains, also considering two transversal dimensions and seven OPs regarding the integrated intervention at the territorial level (Vaquero et al., 2020).

The defined TOs are concentrated on a series of priorities as R&I (TO1), ICT (TO2), business competitiveness (TO3), and the LCE (TO4) that fulfill regulatory requirements (74% in less developed regions; 69% in the region of transition of Algarve; 67% in the outermost regions of Madeira and Azores; 73% in the more



(a) (b)

Fig. 3 ERDF "Total Eligible Costs Decided" and "Total spending" (Eligible Expenditure Declared) for **a** Less Developed Regions, and **b** More Developed, Transition, and outermost regions. *Source* https://cohesiondata.ec.europa.eu/funds/erdf developed region of Lisbon). According to the data available on the P2020 website, about 95% of the budget allocated to Portugal was already spent.

Regarding the evaluation of the PT programs, many difficulties were found regarding data collection. These were due namely to the usage of different reporting nomenclatures across the various PT regions and OPs (see e.g., the reports available at https://portugal2020.pt/portugal-2020/o-que-e-o-portugal-2020/). The autonomous reporting of the lists of Operations referring to P2020 can partially justify those differences (e.g., an ID of investment priority named as 1.1, 1.a, and 1a; the same about indicator CO01 or O.04.02.02.C for the same indicator in investment priority 4.b). Also, these generate some inconsistencies while comparing data at the regional level (i.e., Norte2020, Centro 2020, Alentejo2020, Lisboa2020, Algarve2020 and Madeira e Açores 2020). Another important issue regarding the reporting practices is related to the monitoring periods. The details on data are not uniform for the various programs or the monitored variables. Therefore, making it very hard or even impossible the development of comparative studies.

Concerning the assessment indicators, the EU Observations Report on the PT-funded programs noticed a very positive evolution in Norte OP regarding the construction and definition of the indicator framework. That evolution resulted from (i) the interaction between the regional entities responsible for the programming process and the evaluation team; (ii) the process of harmonizing the bases of indicators at the national level; (iii) the interaction between national entities and the EC (European Commission, 2014a, b).

Following the main funding priorities for the PT programs, our study focuses on three TOs: LCE; R&I; and ICT. To perform the analysis, data were collected from public EU sources on ERDF, the dataset was afterward analyzed, and some preliminary results are illustrated in Fig. 3a and b.

As it can be observed in Fig. 3a, for the "Less Developed Regions", the programs showing higher values are "Competitiveness and Internationalization" (i.e., TO3) followed by the OPs of "Norte" and "Centro" regarding both, the "Total Eligible Costs Decided" and the "Total spending (Eligible Expenditure Declared)" (see the ESIF_2014-2020 categorization file available at https://cohesiondata.ec.europa.eu/funds/erdf).

In 2020, more developed regions, "Lisboa" and "Madeira", registered quite different behaviors, with Madeira having a "Total spending" around half of the "Decided value", while Lisboa registered a lower rate of execution. As "Outermost Regions", Azores and Madeira (i.e., funds 17,740,636,416€ and 11,207,915,405€) both presented an even better execution rate in 2020—see Fig. 3b.

In less developed PT regions, i.e., Norte, Centro, Alentejo, and Azores; the highest "Decided values" were assigned to Multi Thematic Objectives, immediately followed *ex aequo* by "Educational and Vocational Training" and R&I, subsequently followed by the "Competitiveness of Small and Medium Enterprises". The LCE and ICT are in the detached area of the graphic (i.e., split for representing less than 10%), with percentages around 1% or even lower—see Fig. 4.

Concerning the regions of Transition i.e., Algarve; the highest percentage is allocated to "SMEs' Competitiveness" with 43% of the funds. R&I, LCE and ICT

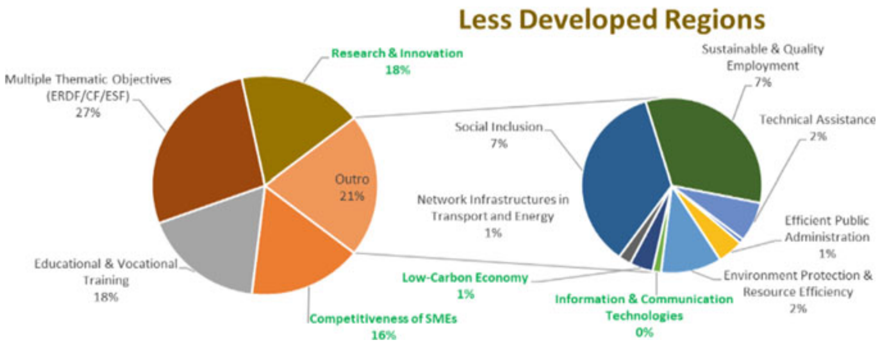


Fig. 4 Total eligible costs decided for less developed regions, 2014–2020, ERDF allocated to PT, evaluated for 2020 as a function of the TOs

have higher representativeness than that observed in less developed regions, but the decided value represents less than 10% of the total, see the detached area of the graph—Fig. 5.

In the more developed regions, the R&I TO appears again with a high representative percentage, followed by the SMEs' competitiveness (TO3)—see Fig. 6. The TOs related to LCE, and ICT registered again a lower percentage of the "Total decided".

For the "Total eligible costs decided" it can be noticed that Multiple Thematic Objectives have the highest values, followed by R&I (TO1) and by Competitiveness of SMEs (TO3), whereas LCE (TO4) and ICT (TO2) obtain the lowest values—see Fig. 7a. Similar behavior is observed for the "Total spending", only with the "Educational and Vocational Training" registering the highest values—see Fig. 7b.

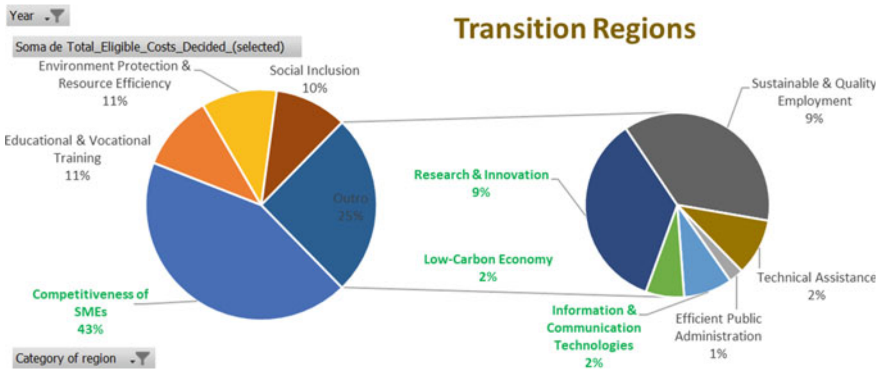


Fig. 5 Total eligible costs decided for transition regions, 2014–2020, ERDF allocated to PT, evaluated for 2020 as a function of the TOs

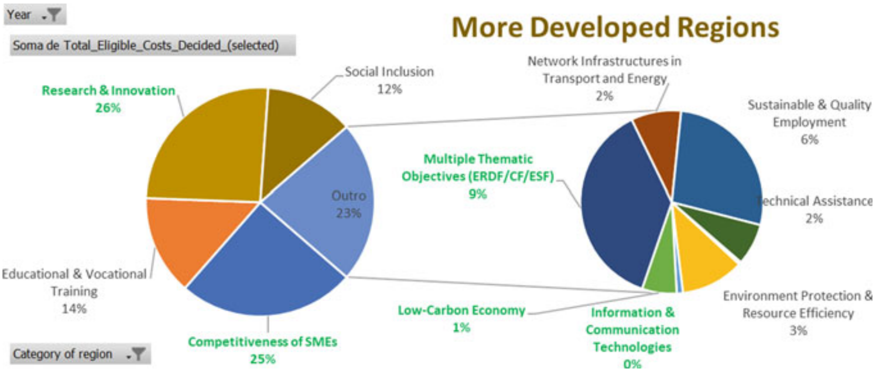
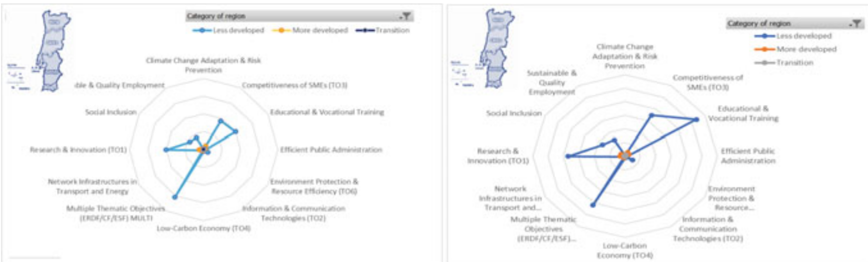


Fig. 6 Total eligible costs decided for more developed regions, 2014–2020, ERDF allocated to PT, evaluated for 2020 as a function of the TOs



(a) (b)

Fig. 7 **a** Total eligible costs decided and **b** total spending (Eligible Expenditure Declared) for 2014–2020. *Source* EC, ESIF, available at https://cohesiondata.ec.europa.eu/2014-2020/ESIF-2014-2020

4 Conclusions

This study explores the present key challenges associated with the evaluation of ERDF committed to three TOs: R&I, LCE, and ICT. A study of the literature on ERDF deployment is done, with a particular emphasis on the PT case, as well as its assessment and reporting processes.

The answers to our research questions are given below.

RQ1: “What are the main areas of concern of the studies devoted to the assessment of ERDF?”

The literature review allowed us to conclude that more recent publications dedicated to the assessment of ERDF show a higher concern over the following themes: “sustainability”, “smart specialization”, “high education” and “regional disparities”. In the case of Portugal, most of the studies focus on “regional development”, also contemplating concerns on “competitiveness”, “higher education”, “innovation” or “rural development”.

RQ2: “What are the main challenges inherent to the selection of the variables/indicators in the assessment of ERDF?”

Overall, there is a lack of data availability that makes it difficult to assess all the OPs targeted to be funded. Besides, there is no full match between the financial data and the corresponding achievements per TO and dimensions of intervention.

We were also able to ascertain that the data available to perform the evaluations differ significantly with the type of TO under scrutiny. In this sense, there is scarce data availability on ICT (TO2) both at the regional and firm levels.

Regarding the PT case, the data analyses allowed us to conclude that there is a significant gap in some of the ERDF initiatives, particularly for the TO2 and TO4, i.e., ICT and LCE. Besides, it was possible to identify a scarcity of data publicly available for the PT OPs, in general, and various reporting conflicts. These issues hampered the possibility of performing deeper analysis, involving, for example, an in-depth comparison across regions’ performance or even enabling productivity analysis.

RQ3: “What are the best practices found in the assessment of ERDF?”

We were able to conclude that the best practices highlight the necessity for indicator harmonization and simplification. A possible way to do this would be the inclusion of all indicators in a single guidance document. It would be desirable to further enhance the quality and comprehensiveness of ICT performance data both at the regional and SMEs levels. Finally, it would be preferable to have more comparable statistics based on the usage of fewer metrics than those established from 2014 to 2020.

RQ4: “What are the main gaps found in the methodologies used to assess ESIF funds?”

Only a small number of assessments employ more consistent methodologies, such as statistical analyses or other mathematical tools. Non-parametric methods, such as DEA, have emerged as a significant quantitative option to the standard methodologies used in comparable circumstances. The primary advantage of employing this mathematical technique is the source of data that it can supply to MA on the inefficiency of the OPs when compared to their peers.

DEA also determines the benchmarks of inefficient OPs, and relevant data on the best procedures to follow to achieve efficiency may be gathered. Nonparametric techniques, such as DEA, may handle multiple evaluation criteria. Additionally, DEA can assist in identifying the primary causes of inefficiency, providing policymakers with helpful information on how to address them. Moreover, the DEA approach is easily adaptable to assess various TOs. This type of analysis is particularly relevant if the programs are already in progress because it enables MA to foresee the impact on the efficiency of future changes in output/input levels. Because more robust techniques are lacking during midterm/terminal evaluations, the use of this nonparametric methodology can be especially advantageous and appropriate, because the current metrics for evaluating the Cohesion Policy can be combined with other methods and contextual factors. This may be accomplished by integrating this type of study with the SFA, for example, enabling us to determine if the inefficient outcomes achieved are mostly due to managerial failings, the contextual environment, or statistical noise.

All in all, our findings emphasize the need for harmonization and simplification of the usage of indicators to evaluate the funded OPs. Besides, an important effort should be placed on the reporting of results to allow for better assessments and to avoid poor outcomes. Finally, we have identified a trade-off between the required detail of the achievements reported and the number of indicators used to support their description, i.e., comprehensiveness versus simplicity.

Acknowledgements This work has been funded by European Regional Development Fund in the framework of Portugal 2020—Programa Operacional Assistência Técnica (POAT 2020), under project POAT-01-6177-FEDER-000044 ADEPT: Avaliação de Políticas de Intervenção Cofinanciadas em Empresas. INESC Coimbra and CeBER are supported by the Portuguese Foundation for Science and Technology funds through Projects UID/MULTI/00308/2020 and UIDB/05037/2020, respectively.

References

Auerbach, A. J., & Gorodnichenko, Y. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy, 4*(2), 1–27. https://doi.org/10.1257/pol.4.2.1

Bachtler, J., & Wren, C. (2006). Evaluation of European Union cohesion policy: Research questions and policy challenges. *Regional Studies, 40*(02), 143–153. https://doi.org/10.1080/00343400600600454

Barca, F., McCann, P., & Rodríguez-Pose, A. (2012). The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science, 52*(1), 134–152. https://doi.org/10.1111/j.1467-9787.2011.00756.x

Cortuk, O., & Guler, M. H. (2015). Disaggregated approach to government spending shocks: A theoretical analysis. *Journal of Economic Policy Reform, 18*(4), 267–292. https://doi.org/10.1080/17487870.2014.951046

Crescenzi, R., & Giua, M. (2016). The EU Cohesion Policy in context: Does a bottom-up approach work in all regions? *Environment and Planning A: Economy and Space*, *48*(11), 2340–2357. https://doi.org/10.1177/0308518X16658291.

Dall'Erba, S., & Le Gallo, J. (2008). Regional convergence and the impact of European structural funds over 1989–1999: A spatial econometric analysis. *Papers in Regional Science, 87*(2), 219–244. https://doi.org/10.1111/j.1435-5957.2008.00184.x

Duranton, G., & Venables, A. J. (2021). Place-based policies: principles and developing country applications. In Handbook of regional science (pp. 1009–1030). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-60723-7_142.

European Commission. (2013). Results indicators 2014? Report on pilot tests in 23 regions/OPs across 15 MS of the EU. Retrieved 30/05/2022. https://ec.europa.eu/regional_policy/sources/docoffic/2014/working/result_indicator_pilot_report.pdf.

European Commission (2014a). Resumo do acordo de parceria para Portugal. 2014, 2014–2020. Bruxelas. Retrieved 30/05/2022. https://ec.europa.eu/info/sites/default/files/partnership-agreement-portugal-summary-july2014_pt.pdf.

European Commission. (2014b). Commission observations—partnership agreement Portugal. Bruxelas. Retrieved 30/05/2022. https://ec.europa.eu/info/sites/default/files/partnership-agreement-portugal-observations-may2014b_en.pdf.

European Commission. (2015). Guidance document on monitoring and evaluation—European regional development fund and Cohesion fund. Concepts and recommendations. Retrieved 30/05/2022, http://ec.europa.eu/regional_policy/sources/docoffic/2014/working/wd_2014_en.pdf.

European Commission. (2018). ERDF & Cohesion fund indicators 2021+, 29 November 2018. https://ec.europa.eu/regional_policy/sources/docgener/evaluation/doc/29112018/6_Indicators 2021_evalnet.pdf.

European Commission. (2022). European structural and investment funds. Country Data for Portugal. Retrieved 01/08/2022. https://cohesiondata.ec.europa.eu/countries/PT.

Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega, 17*(3), 237–250. https://doi.org/10.1016/0305-0483(89)90029-7

Gómez-García, J., Enguix, M. D. R. M., & Gómez-Gallego, J. C. (2012). Estimation of the efficiency of structural funds: a parametric and nonparametric approach. *Applied Economics*, *44*(30), 3935–3954. https://doi.org/10.1080/00036846.2011.583224.

Gouveia, M. C., Henriques, C. O., & Costa, P. (2021). Evaluating the efficiency of structural funds: An application in the competitiveness of SMEs across different EU beneficiary regions. *Omega*, *101*, 102265. https://doi.org/10.1016/j.omega.2020.102265

Gramillano, A., Familiari G., Alessandrini, M., Čekajle, V., Vozab, J., Gassen, N.S., & European Commission (2018b). Development of a system of common indicators for European Regional Development Fund and Cohesion Fund interventions after 2020—Part II—Thematic Objective 2, 7, 8, 9, 10, 11., t33 srl, SWECO consortium, 26th July—2018b. Retrieved 30/05/2022. https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/indic_post2020/indic_post2020_p2_en.pdf.

Gramillano, A., Familiari G., Alessandrini, M., Čekajle, V., Vozab, J., & Gassen, N.S. European Commission. (2018a). Development of a system of common indicators for European Regional Development Fund and Cohesion Fund interventions after 2020—Part II—Thematic Objective 1, 3, 4, 5, 6., t33 srl, SWECO consortium, 26th July—2018a. Retrieved 30/05/2022. https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/indic_post2020/indic_post2020_p1_en.pdf.

Henriques, C., & Viseu, C. (2022a). Are ERDFs Devoted to Boosting ICTs in SMEs Inefficient? Insights through different DEA models. In EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities. In C. Henriques., C. e Viseu, (Eds.), *Proceedings of the 1st International Conference on Evaluating Challenges in the Implementation of EU Cohesion Policy (EvEUCoP 2022)*, (pp. 23–32). Springer.

Henriques, C., & Viseu, C. (2022b). Are ERDFs Devoted to Boosting ICTs in SMEs Inefficient? Further insights through the joint use of DEA with SFA models. In EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities. In C. Henriques, C. e Viseu, (Eds.), *Proceedings of the 1st International Conference on Evaluating Challenges in the Implementation of EU Cohesion Policy (EvEUCoP 2022)*, (pp. 33–43). Springer.

Henriques, C., Viseu, C., Neves, M., Amaro, A., Gouveia, M., & Trigo, A. (2022a). How efficiently does the EU support research and innovation in SMEs? *Journal of Open Innovation: Technology, Market, and Complexity, 8*(2), 92. https://doi.org/10.3390/joitmc8020092

Henriques, C., Viseu, C., Trigo, A., Gouveia, M., & Amaro, A. (2022b). How efficient is the cohesion policy in supporting small and mid-sized enterprises in the transition to a low-carbon economy? *Sustainability, 14*(9), 5317. https://doi.org/10.3390/su14095317

Hollanders, H. (2021). Regional innovation scoreboard 2021. Brussels, Belgium: European Commission. https://doi.org/10.2873/674111.

Kyriacou, A. P., & Roca-Sagalés, O. (2012). The impact of EU structural funds on regional disparities within member states. Environment and Planning C: Government and Policy *30*(2), 267–281. https://doi.org/10.1068/c11140r

Nigohosyan, D., & Vutsova, A. (2018). The 2014–2020 European regional development fund indicators: The incomplete evolution. *Social Indicators Research, 137*(2), 559–577. https://doi.org/10.1007/s11205-017-1610-8

Ortiz, R., & Fernandez, V. (2022). Business perception of obstacles to innovate: Evidence from Chile with pseudo-panel data analysis. *Research in International Business and Finance, 59*, 101563. https://doi.org/10.1016/j.ribaf.2021.101563

Pastor, J., Pons, E., & Serrano, L. (2010). Regional inequality in Spain: permanent income versus current income. *The Annals of Regional Science*, *44*(1), 121–45. https://doi.org/10.1007/s00168-008-0236-9.

Portugal. Portugal 2020—Lista de Operacões Aprovadas. (2020b). Retrieved 23.07.2022. https://www.portugal2020.pt/content/lista-de-operacoes-aprovadas.

Portugal. Webpage P2020a. (2020a). Retrieved 23.07.2022. www.portugal2020.pt.

Puigcerver-Peñalver, M. C. (2007). The impact of structural funds policy on European regions growth. A theoretical and empirical approach. The European Journal of Comparative Economics *4*(2), 179–208. https://doi.org/10.1016/j.giq.2020.101562

Reggi, L., & Gil-Garcia, J. R. (2021). Addressing territorial digital divides through ICT strategies: Are investment decisions consistent with local needs? *Government Information Quarterly, 38*(2), 101562. https://doi.org/10.1016/j.giq.2020.101562

Ruiz-Rodríguez, F., Lucendo-Monedero, A. L., & González-Relaño, R. (2018). Measurement and characterisation of the digital divide of Spanish regions at enterprise level. A comparative analysis with the European context. Telecommunications Policy *42*(3), 187–211. https://doi.org/10.1016/j.telpol.2017.11.007

Scandizzo, P. L., & Pierleoni, M. R. (2020). Short and long-run effects of public investment: Theoretical premises and empirical evidence. *Theoretical Economics Letters, 10*(04), 834. https://doi.org/10.4236/tel.2020.104050

Scotti, F., Flori, A., & Pammolli, F. (2022). The economic impact of structural and Cohesion funds across sectors: Immediate, medium-to-long term effects and spillovers. *Economic Modelling, 111*, 105833. https://doi.org/10.1016/j.econmod.2022.105833

Sörvik, J., & Kleibrink, A. (2016). Mapping EU investments in ICT-description of an online tool and initial observations. Luxembourg: Publications Office of the European Union. Retrieved 30/05/2022. https://op.europa.eu/en/publication-detail/-/publication/6fd75d00-44d0-11e6-9c64-01aa75ed71a1/language-en.

Vaquero, P., Dias, M. F., & Madaleno, M. (2020). Portugal 2020: Improving energy efficiency of public infrastructures and the municipalities' triple bottom line. *Energy Reports, 6*, 423–429. https://doi.org/10.1016/j.egyr.2020.11.195

European Structural and Investment Funds 2021–2027: Prediction Analysis Based on Machine Learning Models

Victor Santos

ABSTRACT This research presents several machine learning algorithms and prediction models to anticipate the European Structural and Investment Funds (ESIF) application in different European Union (EU) countries. These analyses start with data training from 2014 to 2020 ESIF, to test and predict the application of the future ESI Funds for 2021–2027. We deliver an analysis focused on the priorities of each fund, highlighting the differences between the programs in different time periods. In the framework of the European Regional Development Fund (ERDF), we will specifically address the assessment of the following themes: support innovation of small and medium-sized businesses, to greener, low-carbon, and resilient projects with enhanced mobility. In what concerns the European Social Fund (ESF), we will evaluate projects that promote and increase the EU's employment, social, education, and skills policies, including structural reforms in these areas. Regarding the cohesion funds (CF), we will be targeting the improvements between the two ESIFs, looking at projects in the field of environment and trans-European networks in the area of transport infrastructure (TEN-T). In summary, we will be looking at the future of ESIF through the glasses of artificial intelligence.

Keywords European structural and investment funds · Predictive analysis models · Predictive algorithms · CRISP-DM

1 Introduction

The studies available on EU funds have mainly been focused on their implementation by country (Nigohosyan & Vutsova, 2017), by region (Iribas & Pavia, 2010), or even by program (Andrade, 2016). The previous studies focus on the different levels of investments made or on their percentage use in local projects (De Iuliis, 2016). Despite the prolific methods available to perform the evaluation of ESIF, predictive analytics has not yet been applied in this context.

V. Santos (✉)
Polytechnic of Coimbra, Coimbra Business School, ISCAC, Coimbra, Portugal
e-mail: vsantos@iscac.pt

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_11

In this framework, this work aims at proposing a broader and higher perspective regarding the use of predictive analysis. Towards this end, the analysis performed herein is done by trying to predict the total investment that will occur during the implementation of the next programming period of 2021–2027, according to the data available for the previous programming period of 2014–2020. Therefore, the purpose is to predict the total amount that will be devoted to each fund based on the difference between the total investment funds spent and the total amount planned, for each fund, i.e., ERDF, ESF, Youth Employment Initiative (YEI), CF, European Agricultural Fund for Rural Development (EAFRD), and European Maritime and Fisheries Fund (EMFF).

An artificial intelligence tool has been used, specifically the Rapidminer, to test the best prediction model for each ESIF. In this way, it will be possible to anticipate the amount that should be assigned, thus decreasing the difference between the amount spent and planned.

This paper is structured as follows. Section 2 delivers the literature review on the subject. Section 3 briefly goes through the main underpinning assumptions of the methodology employed. Section 4 provides the description of the implementation of the models. The discussion of results is given in Sect. 5. Finally, conclusions are drawn, highlighting the main limitations found and future work developments.

2 Literature Review

Several authors underline the relevance and contribution of big data analytics and the use of machine learning in predictive analytics, reinforcing the role and contribution of decision-making based on business environments (Lismont et al., 2017; Meyer et al., 2019; Psarras, et al., 2020). Quite important also is the methodology and all the processes to achieve the results. Ge et al. (2017), describe all the required processes for extracting the dataset from the database, namely the analysis of the metadata for data preparation and exploration. Only after the dataset identification and preparation, the regression model is selected to perform the analysis. Several hypotheses could be extracted from the application of the different regression methods. In this context, Linear Regression, Support Vector Regression (SVR), Artificial Neural Network (ANN), k-Nearest, Neighbors (k-NN), and M5 model tree are some of the regression models used according to the main goal, which could be employed to model the relationship between the dependent and independent variables. Each model has its own merits and demerits. If the main concern is to maintain the error framed to a short interval, the SVR should be used (Hotzlast, 2022). Although the CRISP-DM is not new, it is a model quite tested and serves as the main structure for the data science process.

3 Methodology

The predictive analytics process described uses the CRISP-DM process model is the most used and common data science process. The step-by-step analysis of the CRISP-DM focuses its attention on the different predictive models. In many cases, it will be the user, not the data scientist, who will carry out the deployment steps. He/she will test the model application for his/her business values (i.e., model hyperparameters). This means that the model should be generic enough for the adaptation to different business variables.

3.1 Rapidminer Automation Procedure

To test the most used predictive methods, the Rapidminer was used as a data science platform that allows data engineering, model building, and machine learning operation, among others. It allows the application of the CRISP-DM model. Therefore, it was used to do the prediction analyses and simulations for each of the ESIFs. This tool has a two-phase automation procedure: the TurboPrep for data preparation and the AutoModel, to test and simulate the different prediction models. For the first 3 phases of the CRISP-DM model, business understanding, data understanding, and data preparation were applied for each of the EU funds, i.e., ERDF, CF, ESF, EAFRD, and EMFF.

Then, after the dataset preparation, the prediction models are simulated and tested using the AutoModel, in order to fulfill all the process modeling phases.

4 Implementation Models

4.1 Data Preparation of EU Funds Using Rapidminer TurboPrep

The first step was to collect each dataset for each fund. The data source was directly obtained from the European Commission data center. The preparation was initiated by identifying every attribute and its meaning in the dataset. This corresponds to step two of the CRISP-DM model, as given below:

Ms—country initials.

Programme Title—program name.

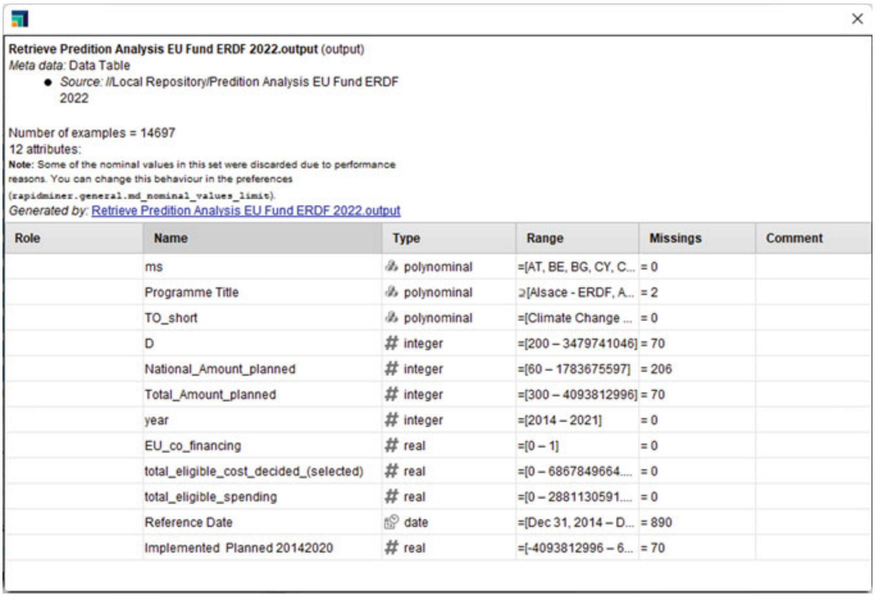
TO_short—main program thematic objective.

National_Amount_planned—investment planned for each country.

Total_Amount_planned—investment planned with all the contributions.

Year—year id of the fund.

EU_co_financing—percentage of European financing over total investment.



Retrieve Predition Analysis EU Fund ERDF 2022.output (output)
Meta data: Data Table

- *Source:* //Local Repository/Predition Analysis EU Fund ERDF 2022

Number of examples = 14697
12 attributes:
Note: Some of the nominal values in this set were discarded due to performance reasons. You can change this behaviour in the preferences (rapidminer.general.md_nominal_values_limit).
Generated by: Retrieve Predition Analysis EU Fund ERDF 2022.output

Role	Name	Type	Range	Missings	Comment
	ms	polynominal	=[AT, BE, BG, CY, C...	= 0	
	Programme Title	polynominal	⊇[Alsace - ERDF, A...	= 2	
	TO_short	polynominal	=[Climate Change ...	= 0	
	D	integer	=[200 – 3479741046]	= 70	
	National_Amount_planned	integer	=[60 – 1783675597]	= 206	
	Total_Amount_planned	integer	=[300 – 4093812996]	= 70	
	year	integer	=[2014 – 2021]	= 0	
	EU_co_financing	real	=[0 – 1]	= 0	
	total_eligible_cost_decided_(selected)	real	=[0 – 6867849664....	= 0	
	total_eligible_spending	real	=[0 – 2881130591....	= 0	
	Reference Date	date	=[Dec 31, 2014 – D...	= 890	
	Implemented Planned 20142020	real	=[-4093812996 – 6...	= 70	

Fig. 1 ERDF MetaData table. *Source* Author's own elaboration

Total_eligible_cost_decided_(selected)—total investment costs of all programs for each country.

Total_eligible_spending—after eligible costs, what was spent.

Reference Date—when the investment was available.

Some attributes were eliminated because of their redundancy, like the country's name and the program's acronym. After that, the metadata analysis table was obtained—see Fig. 1.

From Fig. 2 it can be seen that a new attribute was created for this search, calculated by the difference between the total amount spent and the total amount planned. This attribute "Implemented—Planned 20,142,020" was defined for all the countries during all years between 2014 and 2020.

A summary of the distribution of the new variable can be shown in Fig. 3 and details on the statistics of central tendency are obtainable as well.

4.2 *Data Modeling of EU Funds Using RapidMiner AutoModel*

The next phase of the CRISP-DM involves modeling and simulation for the ERDF. In this framework, the Auto model selects and executes all the predictive models

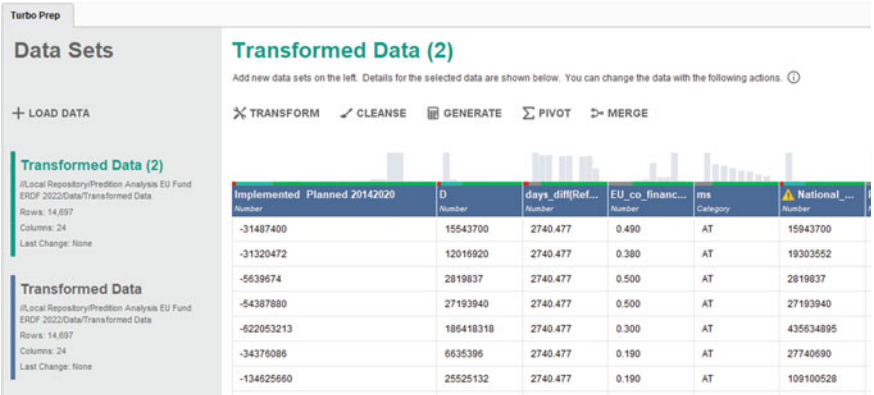


Fig. 2 ERDF data preparation with TurboPrep. *Source* Author's own elaboration

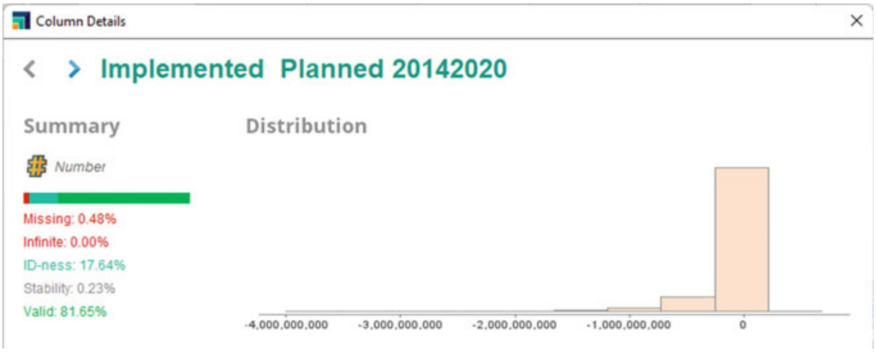


Fig. 3 Measures of central tendency. *Source* Author's own elaboration

available in its library. The dataset of the ERDF with the new attribute was tested with all possible predictive models as shown in Figs. 4, 5 and 6.

The same procedure was done for the other five EU funds applying all the steps of the CRISP-DM model with the TurboPrep for data preparation and exploitation, followed by the modeling analysis on the Automodel.

4.3 Simulation

The Automodel enables choosing the best prediction model, with the best results and lower relative error, by simulating, for each model, the best value for the implementation, which is the difference between the total amount spent and the total amount planned. For the ERDF, the following values were achieved to consider the dependent variables according to Figs. 7, 8 and 9.

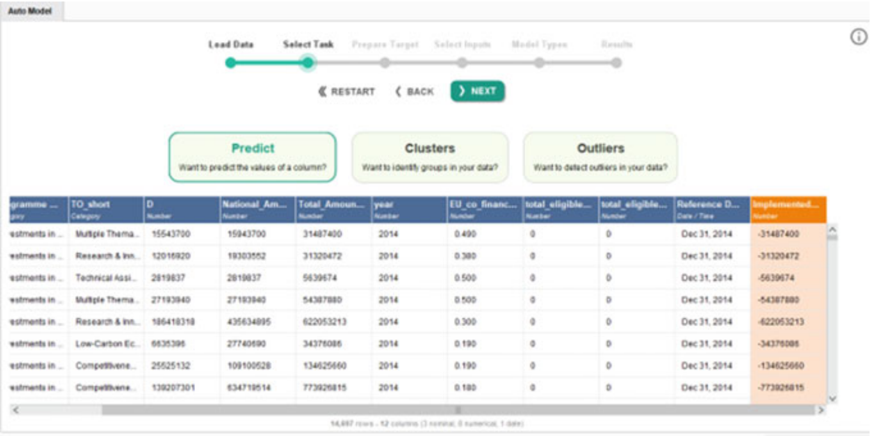


Fig. 4 ERDF Automodel prediction preparation - Part 1. *Source* Author's own elaboration

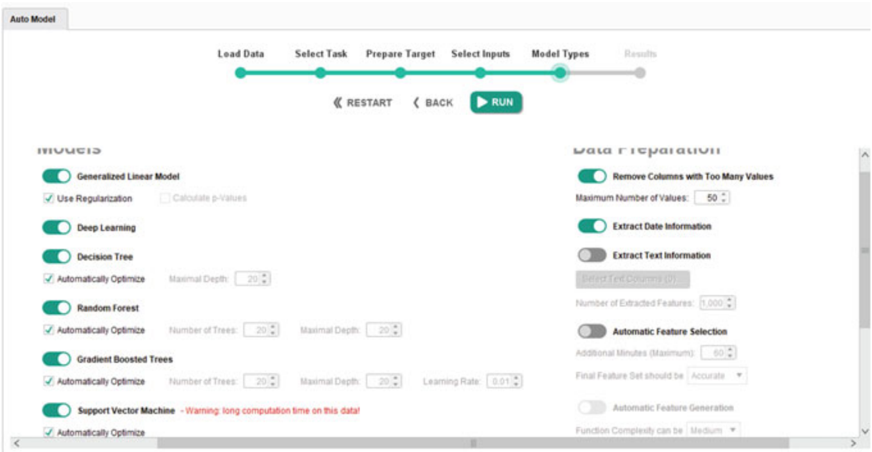


Fig. 5 ERDF Automodel prediction preparation - Part 2. *Source* Author's own elaboration

All the phases of the CRISP-DM model were repeated for the other five EU funds. In the case of the ESF, the YEI, and CF the simulator presented the best prediction model, which was the Generalized Linear Model, while for the remaining funds the best prediction model was the decision tree. From the values previously achieved it is possible to build a prediction table with all the funds, the best prediction algorithm, and their results—see Table 1.

The values presented above (Table 1), validate that there are some funds whose execution differs quite significantly from the amount planned (i.e., present high negative values). Special attention should be given particularly to the execution of

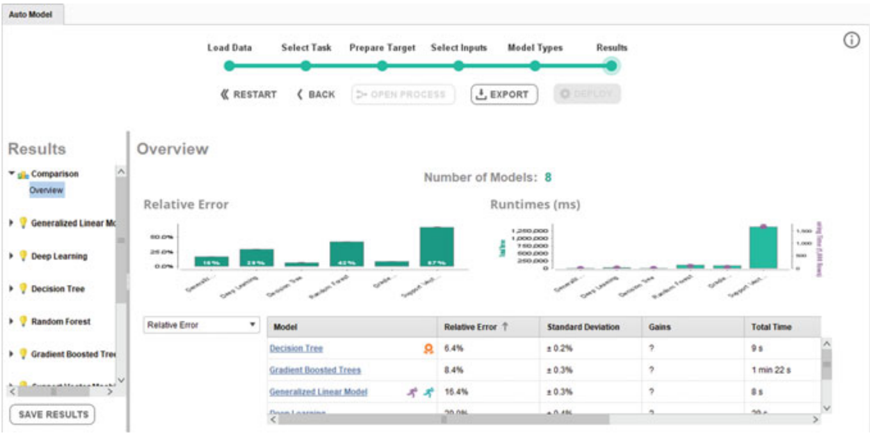


Fig. 6 An overview of the prediction model analysis results for ERDF. *Source* Author's own elaboration

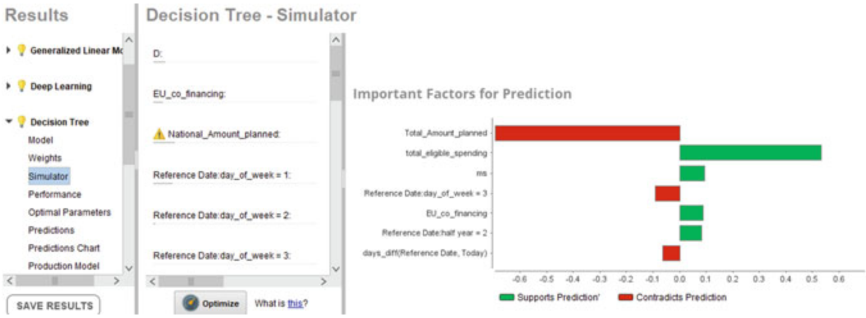


Fig. 7 Important factors for prediction on the decision tree—simulator. *Source* Author's own elaboration

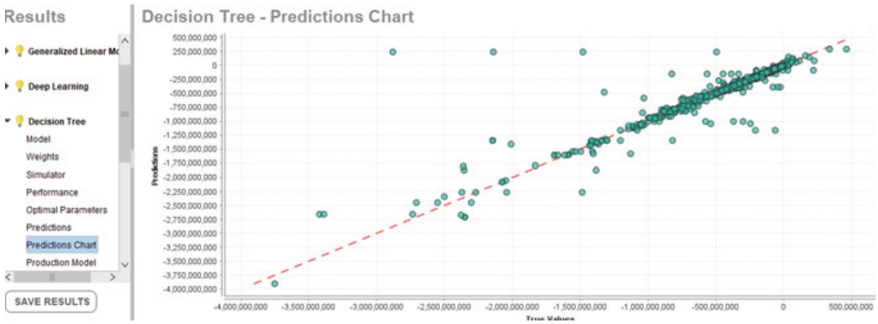


Fig. 8 Decision tree prediction chart. *Source* Author's own elaboration

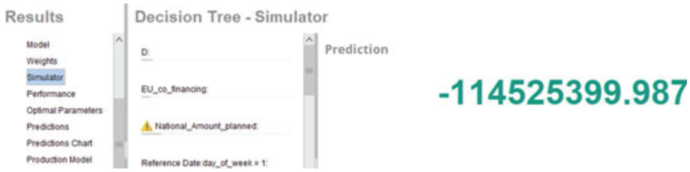


Fig. 9 Final prediction value with the decision tree. *Source* Author's own elaboration

Table 1 Final prediction values and best prediction algorithm for all the ESIF funds from 2014–2020

Fund	Algorithm	Best prediction (€)
ERDF	Decision tree	−114,525,399.98
ESF	Generalized linear model	−99,689,337.84
YEI	Generalized linear model	−155,441,623.41
CF	Generalized linear model	−607,929,889.64
EAFRD	Decision tree	−750,440,879.00
EMFF	Decision tree	−5,635,428,302.54
Total amount allocated 2014–2020		525,729,732,702.00
Total amount allocated 2021–2027 (dated 2021)		215,200,372,408.00

Source Author's own elaboration

EMFF. Although some priorities for the 2021–2027 ESIF are different and have been reduced from eleven to five objectives, this research proposes a method to reduce the total amounts planned.

5 Conclusions and Further Research

By using the CRISP-DM and an artificial intelligence tool like Rapidminer it is possible to achieve some predictions for the next ESIF 2021–2027. Further and deeper research should be taken for each country following the new priorities and using machine learning algorithms for better predictions. A quite interesting study could be developed, in the future with the introduction of new variables, accounting for the impact of the war in Ukraine and the inflation due to the lack of materials, mainly electronic components. Nowadays, these two facts are a concern and will affect the first five priorities and project goals of ESIF. Future work should also involve a similar analysis per each thematic objective.

References

Andrade, P. (2016). Financiación de proyectos culturales con fondos de la Política de Cohesión Europea: Análisis y experiencias en Andalucía 2007–2013.

De Iuliis, C. (2016). The European territorial cooperation. Analysis of results in the seven-year programming period 2007–2013 and the next new programming strategies. *Juridical Current, 19*(4), 85–92. 8p.

Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The role of machine learning. IEEE Access 520590-20616 8051033, https://doi.org/10.1109/access.2017.2756872

Hotzlast, N. (2022). *What is CRISP DM Life Cycle*. Retrieved June 29, 2022, from https://www.datascience-pm.com/crisp-dm-2/

Iribas, B., & Pavia, J. (2010). Classifying regions for European development funding. *European Urban and Regional Studies*, EURO-COMMENTARY.

Lismont, J., Vanthienen, J., Baesens, B., & Lemahieu, W. (2017). Defining analytics maturity indicators: A survey approach. *International Journal of Information Management, 37*(3). S0268401216305655, https://doi.org/10.1016/j.ijinfomgt.2016.12.003.

Meyer, H., Reudenbach, C., Wöllauer, S., & Nauss, T. (2019). Importance of spatial predictor variable selection in machine learning applications–Moving from data reproduction to spatial prediction. *Ecological Modelling*, 411108815-S0304380019303230 108815. https://doi.org/10.1016/j.ecolmodel.2019.108815

Nigohosyan, D., & Vutsova, A. (2017). *The 2014–2020 European Regional Development Fund Indicators: The Incomplete Evolution*. Springer.

Psarras, A., et al. (2020). *Applying the Balanced Scorecard and Predictive Analytics in the Administration of a European Funding Program*. Administrative Science, MDPI.

Technology Serving Justice

Ana Paula Lopes

Abstract In face-to-face interactions, people are constantly providing information through their body movements (Kendon in Body language communication: An international handbook on multimodality in human interaction, pp. 7–27, 2013). Berlin/Boston: De Gruyter Mouton.), in which gestures are included. These kinetic movements transmit two-thirds of what we communicate (Aghayeva in Khazar J Human Soc Sc 53–62, 2011), and ignoring them means disregarding the complexity of the human communication system (Jones and LeBaron in J Commun 52:499–521, 2002). When communicating, humans create signs, and "these signs are made with very many different means (…). They are the expression of the interest of socially formed individuals who, with these signs, realize (…) their meanings" (Kress in Multimodality: A social semiotic approach to contemporary communication. Routledge, p. 10 [2010]). And the way people understand what others mean to transmit can deeply vary. These different interpretations may originate from each person's experience, prejudice, values, and expectations in life. Therefore, the probability of misunderstanding is vast. In the specific context of a forensic interaction, problems of communication misunderstandings can have serious consequences in a suspect's or in a defendant's life. Globally, body movements are not thoroughly considered when it comes to understanding what a suspect or a defendant really wants to declare. However, on some occasions, the correct interpretation of a kinetic movement could contribute to a fairer judicial decision. Through a consistent micro-analysis of interactions, it is possible to create meaning from body movements. The micro-analysis developed by the author showed that body movements can transmit information that had not been verbally uttered. That information has shown to be of great importance in the context of judicial process analysis.

Keywords Communicative process · Face-to-face interactions · Forensic contexts

A. P. Lopes (✉)
Polytechnic of Coimbra, Coimbra Business School | ISCAC, Coimbra, Portugal
e-mail: alopes@iscac.pt

C. Henriques and C. Viseu (eds.), *EU Cohesion Policy Implementation - Evaluation Challenges and Opportunities*, Springer Proceedings in Political Science and International Relations, https://doi.org/10.1007/978-3-031-18161-0_12

1 Introduction

Communication and forensic contexts are closely linked and interrelated. It is impossible to conceive of the idea of forensic contexts without the existence of communicative acts. Forensic contexts are understood as any situation in which issues related to the law arise, encompassing all stages of a judicial process (Granhag & Strömwall, 2004), and including aspects relating to the language used in this process. They involve everyday issues from all kinds of justice and law enforcement institutions—the courts, police, and detention facilities (the courts and criminal police agencies will be the only institutions described, as they were the only ones possible to contact within the scope of this work). These institutions function under the umbrella of the Law but, even so, have different realities and their own ways of working. In other words, and in a broad sense, here we understand forensic contexts as all situations in which language interacts with the legal, judicial and ethical system (www.linguisticaforense.pt). The individuals who move through forensic contexts are also parts of them, such as judges, public prosecutors, lawyers, criminal investigators, and the common citizen, each one performing a distinct function, with an equally different role and focus. This investigation aims to observe how these law enforcement institutions work, namely the courts and police agencies, in relation to the way they process communication, particularly the discursive strategies used, and interrogation methods followed. This study will highlight the way communication functions in the *Polícia Judiciária* (Criminal Police), notwithstanding the real and everyday relevance of the other police agencies that exist in Portugal, due to the importance of this police agency in combatting crime, both at a national and international level. It also aims to present a technological project, which will be developed by the author, following the first steps of this research.

2 The Courts

The rhetoric used by lawyers when they attempt to defend or accuse an individual is well known. What happens in a courtroom will be better understood if it is seen as a story or narrative (Tiersma, 1999). The argumentation strategy used by forensic professionals is not chosen by chance and is appropriate for the intended purpose. Many scholars of judicial language and, specifically, argumentation strategies used in courtrooms, call this use 'narrative' (Cotterill, 2003; Heffer, 2005).

A narrative can be defined as a series of events communicated in order and in a logical sequence, emphasizing its textual organization (Cotterill, 2003). The discourse presented in courts, principally the opening statement, is also characterized and defined as a narrative discourse, with a logical and very specific sequence, passing through the following phases (Heffer, 2005): orientation, problem, evaluation, and resolution. This discourse can often be decisive in a trial (Cotterill, 2003; Heffer, 2005). Lawyers are aware of this and do not choose arguments at random. A well-articulated, convincing discourse, with strong, plausible arguments and words meticulously chosen for the specific context, can very often not only change the course of a trial that was considered almost decided, but even influence the judge's final decision (Heffer, 2005).

Semantic prosody has real and often decisive effects on decision-making in a trial. It consists of an aura of meaning marked by the placement of words spoken to us (Louw, 1993: 157). That is to say, each word occurs alongside other types of words (collocates), which normally have more positive or negative connotations. Often, there are words that collocate with others with a notably negative connotation, which, when used in a specific context like a trial, are capable of influencing a judge to discredit the defendant.

So, the speech used by public prosecutors and lawyers at a trial should not be disregarded in any linguistic analysis of a legal case. Furthermore, the whole narrative and discursive strategy used throughout the course of a trial should be considered in a linguistic analysis of the matter. This is because it is usually through language that legal proceedings test the applicability of generalizations found in the law, generalizations related to specific issues of behavior (Gibbons, 1994). Therefore, trials are linguistic events. So, language is central to the Law, and the Law, as we know it, is not conceivable without language. Many lawyers pride themselves on their mastery of language and view it as a vital skill for professionals in the field (Gibbons, 1994). In this way, it is not only what is written that is important in legal proceedings, as what is verbalized acquires equal importance in a courtroom. Discourse is as indispensable in Law as the written word: during the defendant's interrogation, for example, or the witness statements (Schane, 2006).

3 Criminal Police Agencies

In parallel, we should not disregard the discourse or investigative strategies used by criminal investigators. In the United States and the United Kingdom, for example, the police have and follow well-researched, defined interrogation strategies appropriate to the situation at hand. In these countries, there are even companies that train criminal investigators and teach them strategies to interrogate a suspect more effectively. Much of this training is built on the issues surrounding interactive communication that are the subject of the present study, namely the analysis of kinetic movements made by a suspect when they are being interrogated. The North American company Reid (some of Reid's clients, in addition to many police departments across the United States, include the US Coast Guard, the FBI, and Homeland Security. On their webpage (http://www.reid.com/training_programs/r_interview.html) there is access to a wide range of information about the company itself, as well as the work it does specifically in the area of training on how to conduct a police interrogation, in particular on Reid's nine steps of an interrogation. It has existed for more than half a century and, among other activities, trains professionals in diverse scientific and geographical areas to conduct criminal investigations effectively. The Reid interrogation method follows nine steps: the positive confrontation, theme development, handling denials, overcoming objections, procurement and retention of the suspect's attention, handling a suspect's passive mood, presenting the alternative question, having the suspect orally relate the details of the crime and also elements of oral and written statements (http://www.reid.com/training_programs/r_interview.html). One can see, therefore, that criminal investigators' discourse is not random either, and that all questions raised to suspects, their order, and even the vocabulary used are carefully chosen and have a well-defined purpose.

So far in Portugal, there are no companies or other entities, public or private, collective or individual, that train criminal investigators to carry out in-depth, rigorous, and methodical interrogations on suspects, and that can, on a substantiated scientific basis, interpret diverse types of human communication.

Here, related to the discursive and interrogation strategies used in trials and criminal police agencies, it is also important to observe, on one hand, to what degree the suspects, defendants, and witnesses have equal command over words when called to provide statements and/or testify—compared with that of the judges, lawyers and criminal investigators; and, on the other, in what way these professionals use words in forensic contexts.

4 Communication in a Forensic Context

The concepts of 'power' and 'ideologies' and their relation to the language we all use in our communication take on a particularly relevant role when that communication occurs in forensic contexts. A discursive event is shaped by situations, institutions, and social structures that, in turn, are shaped themselves (Fairclough et al. 2011a, b). Discourse is influenced by social relationships of power and ideology, but these are also influenced by themselves. It is an interdependent relationship that, to understand it, requires an analysis of the discourse in question, but also a consideration and an understanding of the historical and social context in which it was produced.

In the specific case of communication in forensic contexts, it seems important to find out how social power, abuse of power, dominance, and inequality are produced, reproduced, or resisted by texts and by speech (van Dijk, 2001). Similarly, we need to understand what type of relationships exist in these contexts—whether they are opaque or transparent relationships of dominance, discrimination, power, and control manifested through language and the way discourses are conducted. In other words, it is necessary to find out to what point social (and hierarchical) inequalities are expressed, signaled, and legitimized by the use of language in these contexts (Wodak, 2001). Language is the principal means through which institutions create a coherent social reality that frames their sense of identity (Mumby & Clair, 1997). In the same way, institutions—their workers and individuals who interact with them (for example, the public)—are constantly constructed and reconstructed by discursive practices (Mayr, 2008).

Although police interrogation, for example, is a highly regulated discursive form, being structured around legal requirements, its "institutionality" is constructed through the interaction of participants as they negotiate organizational roles (Heydon, 2005). Besides this institutional character, we cannot forget that discourse, whether in a prison room or a courtroom, is not only produced by individuals with a command of legal language. There is also the discourse produced by suspects, defendants, and witnesses (lay people), normally individuals who do not have command of legal language and who produce discourse in an everyday linguistic register. There is, therefore, an unequivocal linguistic disparity, not to mention the issue of discursive power and dominance, that is, who leads the discourse in a trial or police station and how they do it, and consequently who has more power than the rest. The discourse, whether produced by a suspect, a defendant, or the witnesses in a legal case, is controlled by whoever dominates, that is, the judges, public prosecutors, lawyers, and criminal investigators. Individuals participating in the courtroom interaction who do not belong to this sphere of the Law, do not have total control over their oral contributions (Clark, 1996). Witnesses face limitations on how and the time they have to formulate and carry out their actions in real-time. They do not possess any control over when to speak or act. They go to their dedicated position when they are called to testify, they must respond to the questions put to them, without avoiding the subject of the questions. They do not express themselves how they would in other circumstances—in part because they find themselves under enormous restrictions due to the rules imposed—they must stick to the questions they are asked, they were called to testify for one of the parties in that particular case and, therefore, feel obliged to follow a determined line of dialogue (Heffer, 2005).

There is a set of characteristics of face-to-face interactions that are considered fundamental to communication: (1) copresence, (2) visibility, (3) audibility, (4) instantaneity, (5) evanescence, (6) recordlessness, (7) simultaneity, (8) extemporaneity, (9) self-determination and (10) self-expression. By considering each one of these characteristics we can determine that communication in a courtroom, for example, entails considerable complexity (Clark, 1996).

In this regard and knowing that the main parties involved in a trial share the same physical space, we can state that they are co-present (1) and that they can see each other (2) and hear each other (3). However, at a trial, these basic matters of face-to-face interaction can be affected in several ways: the judge may not be able to see the facial expressions of the defendant and witnesses clearly and, frequently, they must ask them to raise their voices. It can also happen that the parties involved are not able to immediately perceive the behavior and actions of each other (4), although the complexity of the communicative act that takes place in a courtroom does not permit everything to be filtered and fixed in the same way for all those present. The oral means of discourse guarantees, normally, that what is said is rapidly forgotten (5) given that it is not.

Shared Understanding and Memories, Expectations, Presuppositions, Stereotypes, Preconceptions, Discrimination

Every individual has a unique and very particular worldview. However, this individual, apart from their uniqueness as a human being, is not an isolated person. The context where they were born, where they grow up, where they will live and their adult life (and these contexts can vary during the course of their life) will mold their personality, their way of thinking, of seeing the world, their ideologies and assumptions. It will, in the same way, create and help to define the stereotypes, preconceptions, and discriminations they will use to make their value judgments concerning others, thereby molding their attitudes and behavior, making them think x or y in relation to a or b. Think about, for example, the ideas argued by Islamic jihad soldiers about individuals outside their belief system—ideas profoundly marked by the sociocultural context they belong to.

So, each one of us is an individual full of preconceived ideas, sharing memories and beliefs with those closest to us. The entirety of our circumstances will influence and shape what we do and say on a day-to-day basis, whether in a social and family context, or a professional context. Van Dijk (1998) addresses these issues related to ideas, beliefs, and values, arguing that they are products of each person's thoughts or a community or group's collective thinking. The ideas and beliefs that each one of us, or a group, possess about a given reality are not limited to being based on the reality that surrounds us or that which we believe to be true or false but are equally products of the judgments we make—whether we think of something as good or bad, pretty or ugly (Van Dijk, 1998). So, our vision of the world that surrounds us is based on partiality and subjectivity, and that which we judge to be right and fair is often in fact not.

The ideologies that guide us function like a system of ideas, belonging to the field of thought and beliefs. They are, in the same way, a social reality, often associated with groups, conflicts, and interests. They can be used to legitimize the dominance and power of one group over another or symbolize social problems. They are, also, linked to language and discourse, as these mechanisms express ideologies to society (Van Dijk, 1998). Ideologies are not just a set of beliefs, but rather beliefs shared and socially accepted by certain groups. These beliefs are acquired and utilized in social contexts, based on the interest of groups or the social relationships between groups (Van Dijk, 1998). As an example: in a courtroom in Portugal, if a defendant belonging to a football supporters' group is presented to a judge and is accused of having provoked disorder, this defendant, as well as the accusation attributed to them (it does not matter here, for now, whether they are innocent or not), will have all the preconceptions, beliefs and ideologies against them that exist surrounding football supporters' groups in Portuguese society—disorderly groups, prone to acts of violence. Moreover, depending on the football club the defendant belongs to—which may be different (and rival) to the one the judge possibly supports in their time—other beliefs and ideologies will certainly be associated. Now, no one is totally objective and impartial, or completely devoid of beliefs, expectations, preconceptions, and ideologies, enabling them to judge a situation without interference. In the case of this example, even unconsciously, this judge will cognitively activate what they think about the defendant's club circumstances, possibly then judging the case in a conditioned manner.

Communication can therefore be based on the relationship between individual and culture and on the knowledge, memories, stereotypes, and preconceptions inherent in each one of us that mold how we communicate and relate to others. It then becomes important to also address face-to-face interaction from a cognitive perspective, in an attempt to find out how our way of thinking influences how we communicate and interact, particularly the way we gesticulate and move the other parts of our body in face-to-face interaction.

5 The Personal and Sociocultural Experience of Those Who Interrogate and Judge

It seems necessary to understand whom the individuals in the 'comfortable position' are, the people judging and evaluating those accused or suspected of committing a crime. The judges and criminal investigators are, above all, human beings with all their inherent peculiarities. They are professionals that have studied and taken an academic and professional path that has enabled them to attain roles in the justice system and criminal investigation. All the education and training attached to these high-level responsibility positions are important and aim to prepare these individuals for roles they will perform as judges or criminal investigators. In the training they receive, both are alerted to issues of the impartiality of justice, their teachers trying to train them not to be prejudiced individuals or biased in their views and convictions. However, we also know that what happens, in reality, is a little different.

The intention here is not to criticize the personality or performance of judges and criminal investigators on a day-to-day basis, but to reiterate that no human being, no matter the training they have received on this subject, is void of convictions, assumptions, preconceptions, choices, and preferences (Anastácio, 2009). Every judge, just like every criminal investigator, is a human being with their own, conditioned circumstances, who grew up in a particular family, with specific convictions, values, and beliefs. They may have (or not) more or less affinity with a particular political party, football club, or religion, and they will have lived through certain experiences (positive and negative) that molded them into the person they are. There is, therefore, a circumstance or worldview that makes that judge and that criminal investigator—that individual in particular—with those characteristics unique and distinguishes them from others.

However, as much as someone tries to remove themselves from all preconceptions and convictions, trying to be as impartial as possible in their opinions and judgments, it is not humanly possible for there to be total impartiality and objectivity. It will be very difficult, for example, for a Roma individual to be presented as the defendant in a Portuguese court and for the judge to not immediately activate all the ideas and preconceptions they have in their mind regarding that ethnic group. Their role as a judge is to analyze the case at hand using the testimonies presented and the evidence (if there is any) shown to them, independently of the race, ethnicity, or creed of the person in front of them. They should analyze all this material and judge the case based on the law in place. Even when the individuals in question belong to the same culture as the judge, there will always be questions that raise problems—religious differences, differences in social status, education, and the way one lives their life and presents themselves, to name a few. Furthermore, the probability of incorrect and prejudiced judgments increases considerably when in front of a judge from another culture.

Milton Bennett (2004) developed a model that addresses the issue of "intercultural sensitivity". Bennett considers that there are people able to interact with other cultures more easily, and some visibly find this type of interaction more problematic. His model, which he called the Developmental Model of Intercultural Sensitivity, DMIS (Bennett, 1986, 2004; Bennett et al., 2003), describes the phenomenon of the sensitivity individuals have or do not have to different cultures and the various stages this phenomenon can encompass.

As people become more interculturally competent, there appears to be a great change in the quality of their communicational experiences, which Bennett calls the transition from ethnocentrism to ethnorelativism. The author used the term ethnocentrism to refer to the experience of the individual's own culture as "central to reality". In other words, the beliefs and behaviors that people acquire in their primary socialization are not questioned—they experience things exactly the way they are. By "ethnorelativism", Bennett intends the opposite of ethnocentrism—the experience of our own beliefs and behaviors as just one possible organization of reality among various others (Bennett, 2004).

This study by Bennett (2004) serves as the basis for an essential question in the present investigation: to what extent are judges and criminal investigators endowed with this "intercultural sensitivity"? Could it be that when they are faced with an individual from another culture, they activate questions in their mind which will allow them to not evaluate the other in a prejudiced way for belonging to another culture, or will exactly the opposite happen? Will they be able to direct their interrogation and assessment, bearing in mind that they are dealing with a person with a different world view to their own, with ideas, beliefs, and values that are culturally marked and different from theirs? And that these do not necessarily make them better or worse? The responses to these questions vary depending on the judge or criminal investigator in question, and the intercultural sensitivity they may have (or not).

Despite the existence of greater or lesser cultural sensitivity, the human condition of prejudice and partiality will very often be a decisive factor present in all evaluations and judgments that are made. So, there exist several aspects that can influence the opinion of a criminal investigator and a magistrate's judgment in court, however impartial these professionals try to be (and we do not question or doubt that the majority indeed try to be so). Here, an essential part of the human being comes in, linked to beliefs, preconceptions, and values, which is the emotional component.

6 The Emotions of Those Who Interrogate and Judge

What is an emotion? This question was formulated exactly like that by William James (1884), as the title of an article he wrote for Mind (Mind is a scientific publication focused on fields such as Psychology and Neuroscience, among other related areas) more than a hundred years ago (Lewis & Haviland-Jones, 2004).

One of the first and most preeminent authors to study emotion was Charles Darwin (1872). Darwin observed and analyzed facial expressions, as well as gestures/body movements, both in humans and animal species, and argued that while gestures/body movements can be related to thoughts, actions, desires, and fantasies among other things, facial expressions relate only to emotions.

Although Darwin maintained the idea of universality in the expression of emotions, above all through the facial expressions exhibited, he recognized that gestures are not universal, but rather socially learned and culturally marked conventions (Darwin, 1872: xxii). However, many of his contemporaries disagreed with his theories of universality relating to the expression of emotions, namely some anthropologists like Margaret Mead (1901–1978, North-American cultural anthropologist.). She believed facial expressions varied from culture to culture, and that the same expressions meant a different emotion depending on the culture in which it was exhibited (Darwin, 1872: xxiii). Later, Ekman (1992) concluded that while there are emotions that are expressed universally, there are also emotions that belong to a certain social/cultural group and that individuals react not only to life phenomena (music, thunder, physical activity) but also when handling interpersonal interactions (Ekman, 1992).

Damásio (1999), in a brief historical perspective, says that emotions were, for a long time mainly after the works by Darwin, James, and Freud (Charles Darwin, William James and Sigmund Freud (XIX century) studied and investigated emotions. Their works in this area are well known. However, in the twentieth century, emotions were set aside in laboratory studies linked to neurological and cognitive studies, and only later given the importance they deserved (Damásio, 1999), viewed as an overly subjective topic and not rational enough for the neurological and cognitive sciences. Once emotions were understood as antipodes of reason, the ultimate quality of the human being, they did not merit being studied (Damásio, 1999). However, in recent years, the neurological and cognitive sciences have begun to address the topic of emotions in more depth, reporting that it did not make sense to place them at the opposite extreme to reason, as emotion and reason are complementary elements of the human being rather than opposites.

The results of investigations carried out in his laboratory showed that emotion is an integral part of reasoning and decision-making processes. The discoveries came from a study of diverse individuals, entirely rational in the way they led their lives until the moment when, as a result of a neurological injury in specific areas of the brain, they lost a certain group of emotions and, at the same time, lost their capacity to make rational decisions (Damásio, 1999).

Damásio also distinguishes emotion from feeling. He suggests the term feeling should be reserved for the mental, private experience of an emotion, while the term emotion should be used to describe the set of responses that an emotion comprises, many of which are publicly observable. An individual cannot observe a feeling in another person, but they can observe a feeling in themselves when they are conscious and aware of their emotional states. In the same way, no one can observe feelings that are not their own, but some aspects of emotions that are at the foundation of these feelings are manifestly observable by others (Damásio, 1999).

Damásio, like Ekman (1992), believes that there are six primary or universal emotions: happiness, sadness, fear, anger, surprise, and dislike. But there are other emotions, called secondary or social emotions, which include shame, jealousy, blame, and pride. There are further background emotions, like well-being or malaise, calmness, and tension (Damásio, 1999). So, according to the author, emotions are complicated sets of neurochemical responses that form a pattern. All emotions play a regulating role that drives, in one way or another, the creation of advantageous circumstances for the organism that manifests the phenomenon. Although learning and culture alter the expression of emotions and redress them in new meanings, emotions are biologically determined processes depending on innate cerebral mechanisms established by a long evolutionary history. Individual history and the fact that culture has a role to play in the formation of some drivers does not negate stereotyping, automation, and the regulatory objective of emotions (Damásio, 1999).

The exact composition and dynamics of emotional responses are formed in everyone through a unique development and environment. However, the evidence suggests that, on the most part, or even totally, emotions result from the long genealogy of evolutionary tuning. They are part of the bioregulatory mechanisms we are born with to prepare us for survival. There are different forms of expression, just like variations in the exact configuration of stimuli that can induce an emotion in different cultures and individuals. However, it is the similarities, not the differences, that are surprising. These similarities make intercultural relationships possible, and enable art, literature, music, and cinema to cross borders so easily (Damásio, 1999).

Darwin (1872) argues that our expressions of emotions are universal and a product of our human evolution (Darwin, 1872: xxii). Even though we all, as a species, evolved in the same way, and therefore everything related to neurological, cognitive, and biological phenomena are similar among human elements belonging to different cultures, it is perhaps more difficult to accept quietly Damásio's idea that suggests intercultural relationships can develop naturally and without conflicts. The author accepts and assumes that there are developmental and cultural elements in each individual that mold how they externalize their emotions. Now, if, and citing the author, "without exception, men and women of all ages, all cultures, all levels of education and economical status have emotions" and if "there is something very particular about the way emotions are linked to complex ideas, values, principles, and judgments that only human beings can have" (Damásio, 1999: 55), it seems evident that the place where we are born, the way we were educated, ultimately, our circumstances and world view will influence and mold the way each individual expresses themselves emotionally. So, on one side we have the purely biological, cognitive, and neurological part that brings us together as beings from the same species, but on the other, we have our sociocultural side, which distances us as beings belonging to cultures and, in a microcosmic sphere, linguistic communities different from each other.

Despite the universal character of the phenomenon of human emotion, argued by Darwin (1872), but opposed by some other researchers (Birdwhistell, 1970; Klineberg, 1940), we need to bear in mind that a display of emotions can vary from culture to culture (or subculture). As a result, the issue of cultural specificity becomes important (Weigand, 2004). The social life of human beings, in the context of a given cultural environment, creates conditions that differ from culture to culture and, therefore, generates specific needs and emotional responses. For example, members of different cultures learn to fear and appreciate different things. This fact is based on the existence of different cultural standards, levels, and behavioral norms that allow, require or prohibit the expression of this or that emotion in a certain way (Weigand, 2004).

It is accepted that an individual makes kinetic movements and expresses emotions, but that they do it within a certain context, a certain situation, and that those realities will permit whoever is interacting with that individual to interpret what they are feeling. The notion of context is, therefore, fundamental, as it does not appear to make sense that speech, kinetic movements, and expression of emotions are interpreted out of context.

We do not communicate or experience emotions in a vacuum, nor can we say what we are feeling based solely on introspection (Armon-Jones, 1986). But we also learn to give meaning to our experience depending on the context, through our social exposure and our cognitive abilities, which allow us to transform our context by the simple fact of interacting with it (Carpendale, 1997). Our emotional experience is linked to specific contexts and has a unique social history and a current cognitive function. Our unique social history includes our immersion in our culture's beliefs, attitudes, and assumptions. All these factors help us learn what it means to feel something and do something with it. The concepts we attribute to emotional experience are full of nuances and meanings depending on the context, including the social roles we occupy, our gender, and our age (Lewis & Haviland-Jones, 2004).

Overall, we all have contextual, cultural, and individually determined perceptions of life and the people we interact with. These are more or less ruled by the values learned throughout our socialization stage in the community—what is considered good and what is considered bad—a situation that makes it difficult, unless one is trained and prepared for this—to recognize that there are different ways of seeing the world. All this becomes hugely important when one intends to question or judge someone who views reality in a way, sometimes, completely different to us, without it being more or less correct, as can happen in forensic contexts.

Some sources of cultural variation are identical to individual sources of variation: in different cultures, like in different individuals, different things function as desirable or undesirable, worthy of approval or worship, or the contrary, appealing or detestable (Ortony et al., 1990).

Therefore, the conduct of each judge and each criminal investigator should be guided by inter and intracultural sensitivity, trying to comprehend that in front of them is a human being, in an inferior and vulnerable position concerning themselves, whom they are judging and evaluating. Also, whether they have committed a crime or not, there is a whole set of cultural and life circumstances that cannot nor should not be disregarded when attempting to sentence their future. Related to this sociocultural component are the kinesic movements that integrate into all our communicational systems and that, as we have seen, should not be ignored in an analysis of a legal process.

7 Kinetic Movements in Interactions in Forensic Contexts

Gestures and other kinetic movements made during a face-to-face interaction in forensic contexts are often not included in analyses of this type of interaction. When they are, the observations on them are for the most part based on empirical experiences, normally without scientific support. A questionnaire (although the questionnaire was sent to various professionals in the legal field, responses were only received from public prosecutors and criminal investigators from the Criminal Police) was therefore designed to find out if and how body movements are taken into account by prosecutors and the Criminal Police investigators during interrogations of defendants, witnesses, and suspects.

From the responses obtained, the position of public prosecutors and criminal police investigators seems clear: not only do the majority (81.8%) feel the need for training in this area of body movement analysis in interactions, but they also think interactions taking place in forensic circumstances should be recorded on video (88.6%) for investigational purposes.

Currently, except for some cases in the United States, the subject of kinetic movements in interactions is not normally taken into account by judicial systems when making decisions in court. Furthermore, so far only the defendants' and witnesses' discourse is recorded and subsequently transcribed, reproducing only that which was verbalized. In countries like England, and also in Portugal, trials are recorded as audio—so there is no access to images—and the quality of recordings is far from ideal to understand clearly what was spoken in the courtroom.

Transcriptions do not specify whom a given discourse is directed toward or who heard it. Based on the principle that they are accurate records of the words that were spoken (although, often, that is not the case), in the majority of cases, they do not supply important clues about the interaction, such as the emphasis given to words/expressions, intonation (whether it is a question or a confirmation), the pauses (which carry their meaning), interruptions (which can be a significant indicator of the speaker's control and their intention), among many other aspects (Shuy, 1996).

As well as this, and whether due to the lack of quality recordings, or a poor interpretation of the real words by the person who did the transcription (and transcribers are not linguists), these transcriptions often do not correspond to the whole truth of what was verbalized. In the case of Portugal, the transcriptions of an audio recording of a trial are only done if they are requested by the interested parties, usually to reopen or continue to investigate a case or when filing an appeal, with the presentation of the allegations.

A transcription implies the conversion of speech into written language. The essential problem is that speech and writing are different means, with different properties. Writing, as we use it every day, is not an effective means of recording speech, as it does not include oral conventions and many oral characteristics (Halliday, 1999). Furthermore, it is practically impossible to precisely record everything that happens in speech, like intonation, breathing, voice quality, accent, pauses, rhythm, and other important aspects. So, the information that is lost is enormously important (Gibbons, 2003).

As a minimum, communication requires three key elements: a producer, a text, and a receiver, in which "producer' and "receiver" are global terms for a set of possible participants (Goffman, 1987), and the "text" can be any type of language (written or spoken) (Heffer, 2005).

Considering, therefore, the process of human communication as it is, how can one believe that, in the analysis of a legal case, essential aspects of the whole process can be neglected: human communication as a whole? Magistrate judges and public prosecutors are qualified in Law and have an understanding of the laws. They are, first and foremost, human beings endowed with common sense and intelligence. However, their academic path will not alert them to the importance of communication as a whole in the context of a trial.

How we communicate—what we verbalize and the co-discursive movements we make—is, therefore, susceptible to problems of interpretation, even when the message being transmitted is in theory truthful and in line with the reality being discussed.

7.1 Forensic Contexts and Analysis of Body Movements

Until now, little has been developed and investigated in the study of gestures and other kinetic movements in an analysis of face-to-face interactions in forensic contexts. Investigators rarely mention body movements and how they relate to speech in interactions taking place in trials, prisons, and police stations (Matoesian, 2010), but some authors have established this relationship (Broaders & Goldin-Meadow, 2010).

The complete and accurate transmission and transcription of information obtained in inquiries and forensic interrogations are very important for the credibility of the justice system. The way questions are asked in inquiries and forensic interrogations influence the responses of those being interrogated (Cotterill, 2003; Heffer, 2005). Equally, speakers make spontaneous co-discursive gestures that can transmit information that was not verbalized (de Ruiter, 2007; McNeill, 1992). In this way, as transcriptions made in forensic contexts only contain discursive elements (Shuy, 1996), as we have seen, a lot of relevant information is ignored (Broaders & Goldin-Meadow, 2010).

There seems, therefore, to be a need for more in-depth material on these contexts, and interdisciplinary knowledge exchange, as ignoring body movements alongside speech in judicial scenarios means not fully comprehending the complexity of the phenomenon being analyzed (Jones & LeBaron, 2002: 512). This need seems justified by the fact that, in the context of the subjects investigated to date, few studies have included an analysis of gestures and other kinetic movements in face-to-face interactions in forensic contexts. In parallel, there have also been few works that relate kinetic movements to interactions in forensic contexts (Matoesian, 2010).

Until now, few investigators have established a relationship between kinetic movements and speech in forensic contexts (Matoesian, 2010: 541). Analyzing acts of communication in a forensic context while ignoring these movements makes for a poorer analysis and eliminates relevant activities in the communicative process (Maynard, 2006: 477). In a courtroom, lawyers direct their gaze to the witnesses when they address them, they mark the rhythm of the speech with hand/arm movements, they show their hand palm facing upwards when they intend to reveal an inconsistency in the testimony given, and witnesses aim their index finger at the defendant, in a pointing gesture (Matoesian, 2010: 542). There are, then, countless kinetic movements that occur alongside speech in any interactive context, but in a forensic context, these can transmit information that may influence the course of the investigation and the judgment in that legal case. As has already been stated, systematically ignoring either speech or kinetic movements in an analysis of face-to-face interaction—as interactions in forensic contexts are—is to set aside vital components of the communicative behavior of human beings, and consequently carry out an incomplete analysis of the whole phenomenon at hand (Jones & LeBaron, 2002).

7.2 The Importance of the Analysis of Body Movements in Forensic Contexts of Interaction

Bearing in mind what has previously been described and highlighted, it seems clear the importance of an analysis of the face-to-face human communication process. When that communication process takes place in forensic contexts (inquiries, police interrogations, among others), and considering, as shown, that two-thirds of the messages that we transmit are passed through body movements (Aghayeva, 2011), to ignore such a significant amount of information means to lose contents that, in the process of analyzing a judicial case, might be vital.

Therefore, it is important that this type of interaction can be legally video recorded (always respecting and updating, if possible, the legislation in force, so that these recordings become possible) so that the interactions can then be analyzed in detail through software (still a project) which allows the upload of the videos, and the recognition and automatic transcription of the speech. By doing so, it is then possible to reach adequately supported findings about the individuals' communicative behavior in those contexts. It cannot be emphasized enough how important it is to keep these recordings exclusively under the scope of helping in the criminal investigation processes, being confidential material that must be carefully protected.

Because, in fact, human beings can change, omit, and forge information through speech, but they cannot do it permanently through their bodies. It is of major interest to all that justice can comply with its duty, and it is defended here that the interdisciplinarity between Law, Technology, and Linguistics seems to undoubtedly create advantages in ascertaining the truth in a judicial process.

8 Conclusions and Further Research

As such and considering the opinions of the public prosecutors and criminal investigators questioned in the context of the present work, of whom a large majority believe more training in the analysis of interactive body movements is necessary, the argument for the importance of this training gains more traction. Also, implicitly, the need arises for interactions in forensic contexts to be recorded on video—at least for investigational purposes—for a detailed, substantiated, and credible analysis of all information transmitted during interactions to be possible. In a related investigation carried out by the author, it was possible to verify the following—by highlighting just a few aspects—and through an analysis of body movements during interactions:

Through gestures, speakers can transmit information they have not verbalized. In this way, access can be gained to messages and/or mental images that speakers may wish to omit voluntarily or involuntarily, which may turn out to be important depending on the context of the interaction. Gestures thereby function as a window into the mind (de Ruiter, 2007; McNeill, 1992). In addition, pauses and hesitations can reveal that a speaker is organizing their thoughts or, if they are interacting in their second language, that they need more time than they would need in their first language to choose lexical elements or structure the syntax of their discourse.

Common aspects shared by speakers of an interaction (culture, assumptions, expectations, beliefs, ideologies, education…) can help make the interaction more natural and spontaneous, with the message transmitted more effectively and more easily understood.

We confirm, therefore, the importance of understanding and analyzing the communicative process, as not considering two-thirds of the information passed on, particularly for interactions in forensic contexts, means ignoring rich material for criminal investigation.

To try to avoid this loss, and misunderstandings and misinterpretation in verdicts, we think that it is important to include this type of analysis in the wider judicial cases analysis. Therefore, and to help do this faster and in a reasoned and reliable way, the author aims to develop software that, as previously mentioned, will allow the fulfillment of this analysis.

References

Aghayeva, K. (2011). Different aspects of nonverbal intercultural communication. *Khazar Journal of Humanities and Social Sciences*, 53–62.

Anastácio, M. (2009). *Psicologia das Motivações Ajurídicas do Sentenciar: O Lado Invisível da Decisão.* Universidade Lusófona.

Armon-Jones, C. (1986). The social functions of emotion. In R. Harré (Ed.), *The social construction of emotions* (pp. 57–82). Basil Blackwell.

Bennett, J. M., Bennett, M. J., & Allen, W. (2003). Developing intercultural competence in the language classroom. In D. L. Lange & R. M. Paige (Eds.), *Culture as the core: Perspectives on culture in second language learning* (pp. 237–270). Information Age Publishing.

Bennett, M. J. (1986). A developmental approach to training for intercultural sensitivity. *International Journal of Intercultural Relations, 10*(2), 179–196.

Bennett, M. J. (2004). Becoming interculturally competent. In J. S. Wurzel (Ed.), *Toward Multiculturalism: A reader in multicultural education* (pp. 62–77). Intercultural Resource Corporation.

Birdwhistell, R. L. (1970). *Kinesics and context*. University of Pennsylvania Press.

Broaders, S. C., & Goldin-Meadow, S. (2010). Truth is at Hand: How Gesture Adds Information During Investigative Interviews. *Psychological Science*, *21*(5), 623–628.

Carpendale, J. I. (1997). An explication of Piaget's constructivism: Implications for social cognitive development. In S. Hala (Ed.), *The development of social cognition* (pp. 35–64). University College London Press.

Clark, H. H. (1996). *Using Language*. Cambridge University Press.

Cotterill, J. (2003). *Language and power in court: a linguistic analysis of the O.J. Simpson trial.* Palgrave Macmillan.

Damásio, A. (1999). *O Sentimento de Si: O Corpo, a Emoção e a Neurobiologia da Consciência.* Publicações Europa-América.

Darwin, C. (1872). *The expression of emotions in man and animals.* Oxford University Press.

de Ruiter, J. P. (2007). Postcards from the mind: The relationship between speech, imagistic gesture and thought. *Gesture, 7*(1), 21–38.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3–4), 169–200.

Fairclough, N., Mulderrig, J., & Wodak, R. (2011a). Critical Discourse Analysis. In T. Van Dijk, (Ed.) *Discourse Studies: A multidiciplinary introduction.* London: Sage.

Fairclough, N., Mulderrig, J., & Wodak, R. (2011b). Critical Discourse Analysis. In T. Van Dijk (Ed.), *Discourse Studies: A multidisciplinary introduction* (pp. 357–378). Sage.

Gibbons, J. (1994). *Language and the law*. Longman.

Gibbons, J. (2003). *Forensic linguistics: An introduction to language in the justice system.* Blackwell Publishing.

Goffman, E. (1987). *The presentation of self in everyday life.* Penguin.

Granhag, P. A., & Strömwall, L. A. (2004). *The detection of deception in forensic contexts.* Cambridge University Press.

Halliday, M. A. (1999). The notion of "context" in language education. In M. Ghadessy (Ed.), *Text and context in functional linguistics* (pp. 1–24). John Benjamins Publishing Company.

Heffer, C. (2005). *The language of jury trial: A corpus-aided analysis of legal-lay discourse.* Palgrave Macmillan.

Heydon, G. (2005). *The language of police interviewing: A critical analysis.* Palgrave Macmillan.

James, W. (1884). What is an emotion. *Mind, 9*, 188–205.

Jones, S., & LeBaron, C. D. (2002). Research on the relationship between verbal and non-verbal communication: Emerging integrations. *Journal of Communication, 52*(3), 499–521.

Kendon, A. (2013). Exploring the utterance roles of visible bodily action: a personal account. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf, (Eds.), *Body language—communication: An international handbook on multimodality in human interaction* (vol. 1, pp. 7–27). De Gruyter Mouton.

Klineberg, O. (1940). Emotional expression in Chinese literature. *Journal of Abnormal and Social Psychology, 33*, 517–520.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication.* Routledge.

Lewis, M., & Haviland-Jones, J. M. (2004). *Handbook of emotions.* The Guilford Press.

Louw, B. (1993). Irony in the text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli, Text and technology: in honour of john sinclair (pp. 157–176). John Benjamins.

Matoesian, G. M. (2010). Multimodality and forensic linguistics: Multimodal aspects of victim's narrative in direct examination. In M. Coulthard & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (pp. 541–557). Routledge.

Maynard, D. W. (2006). Comment - bad news and good news: losing vs. finding the phenomenon in legal settings. *Law and Social Inquiry, 31 (2)*, 477–497.

Mayr, A. (2008). *Language and power: An introduction to institutional discourse*. Continuum.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago University Press.

Mumby, D. K., & Clair, R. P. (1997). Organizational discourse. In T. Van Dijk (Ed.), *Discourse studies, Volume 2: Discourse as social interaction* (pp. 181–205). Sage.

Ortony, A., Gerald, L. C., & Collins, A. (1990). *The Cognitive structure of emotions.* Cambridge University Press.

Schane, S. A. (2006). *Language and the law*. Continuum.

Shuy, R. W. (1996). *Language crimes: Use and abuse of language evidence in the courtroom.* Blackwell.

Tiersma, P. (1999). *Legal language*. University of Chicago Press.

Van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach.* SAGE Publications Ltd.

Van Dijk, T. (2001). Critical discourse analysis. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 352–371). Blackwell.

Weigand, E. (2004). *Emotion in dialogic interaction.* John Benjamins Publishing Company.

Wodak, R. (2001). What CDA is about—a summary of its history, important concepts and its developments. In R. Wodak & M. Meyer (Eds.), *Methods of Critical Discourse Analysis* (pp. 1–14). Sage.